

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/108633>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

XENOPUS LAEVIS OCTAMER-BINDING PROTEINS.

A thesis submitted for the degree of doctor of philosophy.

Darrin P. Smith.

Animal Molecular Genetics Group,
Department of Biological Sciences,
University of Warwick.

September, 1990.

THE BRITISH LIBRARY DOCUMENT SUPPLY CENTRE

BRITISH THESES N O T I C E

The quality of this reproduction is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print, especially if the original pages were poorly produced or if the university sent us an inferior copy.

Previously copyrighted materials (journal articles, published texts, etc.) are not filmed.

Reproduction of this thesis, other than as permitted under the United Kingdom Copyright Designs and Patents Act 1988, or under specific agreement with the copyright holder, is prohibited.

THIS THESIS HAS BEEN MICROFILMED EXACTLY AS RECEIVED

THE BRITISH LIBRARY
DOCUMENT SUPPLY CENTRE
Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
United Kingdom

CONTENTS.

Table of contents.	ii
Table of figures.	xv
Acknowledgements.	xx
Declaration.	xxi
Summary.	xxii
Abbreviations.	xxiii

Table of contents.

Chapter 1: INTRODUCTION.

1.1	The octamer paradox.	1
1.1.1	The octamer motif and immunoglobulin gene expression.	3
1.1.2	The octamer motif and histone H2B gene expression.	8
1.1.3	The octamer motif and snRNA gene expression.	13
1.1.4	The SV40 enhancer contains an octamer motif.	15
1.1.5	The octamer motif and the expression of HSV alpha (immediate early) genes.	17
1.1.6	The octamer motif and the expression of other genes.	18
1.1.7	The octamer motif and adenovirus replication.	20
1.2	Models of transcriptional activation by Oct-1/Oct-2 from transfection studies in B and non-lymphoid cells.	21
1.2.1	Oct-1 and Oct-2 are transcriptional activators.	22
1.2.2	Transfection of SV40 enhancer-containing constructs into B and non-lymphoid cells.	24
1.2.3	VP16 and the activation of U snRNA genes.	28
1.2.3	VP16 and the activation of H2B transcription.	31

1.3	Studies on cloned Oct-1 and Oct-2A/Oct-2B.	31
1.3.1	Cloning of Oct-1 and Oct-2A/Oct-2B.	32
1.3.2	Oct-1 and Oct-2A/B have a conserved POU domain which directs DNA binding.	34
1.3.3	Other features determined by examination of the sequences of cloned Oct-1 and Oct-2A/B.	38
1.3.4	Transcriptional activation by cloned Oct-1 and Oct-2A.	40
1.3.5	The Oct-1 POU domain interacts with VP16.	44
1.3.6	The POU domain is sufficient to stimulate DNA replication.	46
1.4	Other POU domain and octamer-binding proteins.	47
1.4.1	Hoxo box-containing genes form multi-gene families in <u>Drosophila</u> and vertebrates.	47
1.4.2	The <u>unc-86</u> gene of <u>C. elegans</u> .	49
1.4.3	The pituitary-specific transcription factor, Pit-1.	49
1.4.3	A family of POU domain genes in the mammalian brain.	54
1.4.5	A family of octamer-binding proteins in mouse embryogenesis.	55
1.4.6	Other examples.	58
1.5	Project aims.	60

Chapter 2: MATERIALS.

2.1 General reagents.	61
2.2 Stock solutions.	62
2.3 Bacteriological media.	63
2.4 Bacteria, plasmids and phage.	
2.4.1 Genotypes of <u>E. coli</u> strains.	64
2.4.2 Plasmid vectors.	64
2.4.3 Bacteriophage vectors.	65
2.4.4 Plasmid and Bacteriophage Recombinants.	65

Chapter 3: METHODS.

3.1 Tissue culture.	
3.1.1 Growing <u>X. laevis</u> Xtc cells.	67
3.1.2 Growing mouse L cells.	67
3.1.3 Blocking <u>X. laevis</u> Xtc cells at stages of the cell cycle.	67
3.2 Obtaining oocytes, eggs and embryos.	
3.2.1 Oocytes.	68
3.2.2 Eggs and embryos.	68
3.3 Microinjection of <u>Xenopus</u> oocytes.	69

3.4	Preparation of RNA.	
3.4.1	From <u>X. laevis</u> oocytes.	70
3.4.2	From adult <u>X. laevis</u> liver.	71
3.4.3	From microinjected oocytes.	72
3.4.4	Selection of polyadenylated RNA by oligo dT cellulose chromatography.	72
3.5	Gels for resolving nucleic acids.	
3.5.1	Non-denaturing agarose gels.	73
3.5.2	Low melting point agarose gels.	73
3.5.3	Formaldehyde-agarose RNA gels.	74
3.5.4	Denaturing polyacrylamide gels.	75
3.6	Isolation of genomic DNA from adult <u>Xenopus</u> blood.	76
3.7	Southern blotting.	77
3.8	Northern blotting.	79
3.9	Large scale preparation of plasmid DNA and purification by caesium chloride/ethidium bromide centrifugation.	80
3.10	Primer extension analysis of RNA.	82

3.11 RNase protection assays.	83
3.12 Subcloning techniques.	
3.12.1 Restriction enzyme digests.	84
3.12.2 Preparation of plasmid vectors for subcloning.	84
3.12.3 Preparation of target DNA for subcloning.	85
3.12.4 Ligations.	85
3.13 Transformation of <u>E. coli</u> with plasmids.	
3.13.1 CaCl_2 mediated transformation.	85
3.13.2 Hanahan high efficiency transformation.	86
3.14 Plating <u>E. coli</u> transformed or infected with bacteriophage M13.	88
3.15 Small scale isolation of plasmid DNA and M13 RF DNA.	88
3.16 Preparation of single stranded M13 template DNA.	90
3.17 Complementation tests on single stranded M13 DNA.	91
3.18 Plating bacteriophage lambda.	91
3.19 Bacteriophage lambda plaque lifts.	92

3.20 Automatic excision of phagemids from lambda ZAP clones.	94
3.21 DNA sequencing by the dideoxy chain termination method.	
3.21.1 Sequencing single stranded M13 templates.	95
3.21.2 Sequencing plasmid DNA.	96
3.22 Oligonucleotide directed mutagenesis.	
3.22.1 Phosphorylating synthetic oligonucleotides.	97
3.22.2 Annealing oligonucleotide to template.	98
3.22.3 Extension reactions.	98
3.22.4 Screening for mutants.	99
3.23 SDS-polyacrylamide gel electrophoresis.	
3.23.1 Running protein gels.	101
3.23.2 Staining protein gels with Coomassie blue.	101
3.23.3 Fluorography of protein gels.	102
3.24 Western blots.	102
3.25 Immunodetection of proteins on Western blots.	103

3.26	Generation and purification of a polyclonal antiserum.	
3.26.1	Purification of a fusion protein produced in <u>E. coli</u> for use as an antigen.	104
3.26.2	Immunisation of rabbit.	104
3.26.3	Affinity purification of antiserum.	105
3.27	Sucrose density gradient centrifugation.	106
3.28	Preparation of nuclear extract from Xtc cells.	107
3.29	Methods for radiolabelling DNA and RNA.	
3.29.1	Nick translation.	108
3.29.2	End labelling DNA with γ - ^{32}P -ATP and T4 polynucleotide kinase.	109
3.29.3	Labelling DNA fragments by end-filling.	110
3.29.4	^{32}P labelled synthetic RNA probes.	110
3.30	Preparation of synthetic RNA for microinjection into oocytes.	112

3.31	Preparation of protein extract for band shift assays.	
3.31.1	From <u>X. laevis</u> oocytes, eggs and embryos.	113
3.31.2	From adult <u>X. laevis</u> tissues.	114
3.31.3	From cultured cells.	114
3.31.4	From <u>E. coli</u> expressing fusion proteins.	115
3.32	Band shift assays.	
3.32.1	Preparation of probes.	115
3.32.2	Assays.	115

RESULTS AND DISCUSSION.

Chapter 4: Homologues of human Oct-1 occur in the
X. laevis genome, and in X. laevis
 oocyte RNA.

Introduction	117
4.1 <u>X. laevis</u> genomic Southern blot.	117
4.2 <u>X. laevis</u> A+ selected RNA Northern blot.	118

Chapter 5: Isolation of Xenopus laevis Oct-1 cDNA
 clones from an oocyte library.

5.1 Library screening.	120
5.2 Automatic excision of prospective Oct-1 cDNA clones from the phage vector and preliminary restriction analysis.	121
5.3 Detailed restriction analysis of clones 3, 6 and 16.	121

Chapter 6: Sequence analysis of Oct-1 homologues.

6.1 Sequencing clones 3 and 16.	125
6.2 Comparison of <u>X. laevis</u> Oct-1 protein sequences with human and chicken Oct-1.	127

Chapter 7: Southern blot of X. laevis genomic DNA
 probed with POU domain, Oct-1 specific
 and Oct-2 specific probes.

130

Chapter 8: Detection of Xl-Oct-1A and Xl-Oct-1B transcripts in <u>X. laevis</u> oocyte RNA by RNase protection assays.	
Introduction.	132
8.1 Subcloning of a fragment of Xl-Oct-1A into a transcription vector, for use in the preparation of synthetic RNA probes.	133
8.2 The probe detects both Xl-Oct-1A and Xl-Oct-1B synthetic RNA.	133
8.3 Xl-Oct-1A and Xl-Oct-1B transcripts can be detected in <u>X. laevis</u> oocytes.	134
Chapter 9: Expression of Xl-Oct-1A synthetic RNA in micro-injected oocytes.	136
Chapter 10: Preparation of Xl-Oct-1A fusion protein constructs.	
Introduction.	139
10.1 Making a <u>Bgl</u> II site at the 5' end of Xl-Oct-1A.	140
10.2 Making fusion protein constructs.	141

Chapter 11: Production of an anti Xl-Oct-1A polyclonal antiserum.	
Introduction.	143
11.1 Making an Xl-Oct-1A fusion protein for use as an antigen.	143
11.2 Production of a polyclonal antiserum.	145
Chapter 12: The <u>Xenopus laevis</u> Oct-1 cDNA encodes a functional octamer-binding protein.	147
Chapter 13: Two octamer-binding proteins can be detected in <u>Xenopus laevis</u> oocyte extract.	
13.1 Oocytes contain Oct-1 and a second octamer-binding protein.	150
13.2 An antibody against full-length human Oct-1 recognises <u>Xenopus laevis</u> Oct-1, but not Oct-R.	151
Chapter 14: A comparison of the binding properties and distribution of Oct-1 and Oct-R.	
14.1 Oct-1 and Oct-R have different binding affinities.	153
14.2 Distribution of Oct-1 and Oct-R in tissues and early development.	155
14.3 Oct-R cannot be detected in mouse L cells.	156
14.4 Location of Oct-1 and Oct-R in the oocyte.	156

14.5 Levels of Oct-1 and Oct-R in cells in which DNA synthesis has been inhibited and in cells which have been serum starved.	157
14.6 Conclusions and speculation regarding <u>Xenopus laevis</u> Oct-1 and Oct-R based on the affinity/distribution data.	158
Chapter 15: Do the octamer motif and Oct-1/Oct-R regulate <u>Xenopus laevis</u> histone H2B genes?	
Introduction.	171
15.1 Preparation of a H2A-H2B expression construct.	172
15.2 Expression of pH2A/B.exp can be detected in microinjected oocytes.	175
15.3 An attempt to modulate pH2A/B.exp expression by competition with oct factor binding sites.	175
15.4 Making a mutation in the octamer motif associated with the H2B gene of pH2A/B.exp.	176
15.5 Effect of the octamer mutation on H2A/B expression.	177
Chapter 16: The anti Oct-1 polyclonal antiserum detects proteins other than Oct-1 on Western blots.	
16.1 The anti Oct-1 antiserum specifically detects two proteins in ovary protein extract.	180

16.2 Oct-1 and the proteins to which the anti Oct-1 antiserum reacts on a Western blot can be separated on a sucrose gradient.	182
16.3 Distribution of Oct-1 related proteins in oogenesis and early development.	183
16.4 Location of Oct-1 related proteins in the cell.	184
Chapter 17: General discussion and conclusions.	186
REFERENCES.	190

Table of figures.

Figure number.	Title.	Prior to page.
1	The structure of histone H2B gene promoters.	10
2	A sequence comparison of H2B boxes.	11
3	A model for transcriptional activation by Oct-1 and Oct-2.	26
4	The POU domain.	34
5	Oct-1 binds to degenerate octamer motifs by association with flanking sequences.	37
6	Homology between Oct-1 and Oct-2 outside the POU domain.	38
7	Southern blot of <u>X. laevis</u> genomic DNA probed with human Oct-1 cDNA.	118
8	Northern blot of <u>X. laevis</u> RNA probed with human Oct-1 cDNA.	119

9	A third round screen of a positive cDNA clone from the <u>X. laevis</u> oocyte cDNA library.	121
10	Partial restriction maps of clones 3, 6 and 16 aligned to indicate conservation of some sites.	123
11	Clone 6 does not hybridise to a fragment from the 5' end of human Oct-1.	124
12	Sequencing strategy for clone 3.	126
13	Sequencing strategy for clone 16.	126
14	Complete nucleotide sequence of clone 3 (X1-Oct-1A).	127
15	Complete nucleotide sequence of clone 16 (X1-Oct-1B).	127
16	Nucleotide sequence alignment of clone3 (X1-Oct-1A) and clone 16 (X1-Oct-1B).	127
17	Predicted amino acid sequence alignment of human Oct-1, chicken Oct-1, X1-Oct-1A and X1-Oct-1B.	128

18	Alternate splicing occurs at the N terminal end of human Oct-1.	129
19	Southern blot of <u>X. laevis</u> genomic DNA probed with POU domain, Oct-1 specific and Oct-2 specific probes.	131
20	Detection of Xl-Oct-1A and Xl-Oct-1B transcripts in oocytes by RNase protection assays.	134
21	Expression of Xl-Oct-1A synthetic transcripts in micro-injected oocytes.	137
22	Xl-Oct-1A sequence contained in fusion protein constructs.	141
23	Expression of fusion proteins suitable for use as antigens.	144
24	The anti Oct-1 antiserum detects the <u>E. coli</u> fusion protein against which it was raised.	146
25	Oct-1 sequence contained in fusion protein constructs and their binding to an octamer motif analysed by band shift assays.	148

26	Band shift assays showing that oocyte extract contains Oct-1 and a second octamer-binding protein, Oct-R.	151
27	Anti human Oct-1 antiserum reacts with <u>Xenopus laevis</u> Oct-1, but not Oct-R.	152
28	Comparison of the binding affinities of Oct-1 and Oct-R for different octamer-containing oligos using band shift assays.	154
29	Distributions of Oct-1 and Oct-R determined by band shift assays.	155
30	Band shift assays to show that Oct-R cannot be detected in mouse L cell extract.	156
31	Location of Oct-1 and Oct-R in <u>Xenopus laevis</u> oocytes determined by band shift assays.	157
32	Band shift assays to determine the levels of Oct-1 and Oct-R in <u>Xenopus laevis</u> Xtc cells treated with hydroxyurea or serum-starved.	158
33	<u>X. laevis</u> H2B boxes and band shift probes.	159

34	X1HW23 intergenic region and structure of pH2A/B.exp.	173
35	pH2A/B.exp is expressed in micro-injected oocytes.	176
36	Effect of competition with oct factor binding sites on pH2A/B.exp expression in oocytes.	176
37	Mutagenesis of the octamer motif associated with the H2B gene of pH2A/B.exp.	177
38	A comparison of pH2A/B.exp and pH2B.exp-mut expression in oocytes.	178
39	Anti Oct-1 antiserum specifically detects two proteins in oocyte extract.	181
40	The protein to which the anti Oct-1 antiserum reacts and Oct-1 can be separated on a sucrose gradient.	183
41	Distribution of Oct-1 related proteins in oogenesis and early development.	184
42	Location of Oct-1 related proteins in the cell.	185

Acknowledgements.

I am grateful for the assistance of my supervisor, Bob Old, throughout the course of the project. Also, I would like to thank Glen Sweeney for advice and encouragement.

My thanks are also due to Marcela Vlad for assistance with microscopy, to Liz Jones for help in the preparation of a polyclonal antibody, and to Winship Herr for provision of materials and unpublished data.

I acknowledge R. W. Old for provision of the plasmid pRW23 and an anti X. laevis THR antiserum, W. Herr for provision of the plasmid pBS-Oct-1+ and an anti human Oct-1 antiserum, P. Matthias for provision of the human Oct-2A cDNA, A. R. Brooks for provision of the plasmid albino Δ -670, and J. Shuttleworth for provision of an oocyte cDNA library.

Financial support was provided by the Medical Research Council.

Declaration.

All the results presented in this thesis were obtained by the author, apart from those that are specifically indicated in the text. All the oocyte injections were performed by Bob Old.

All sources of information have been acknowledged by means of reference. None of the work contained in this thesis has been used for any previous application for a degree.

Summary.

The ubiquitous human octamer-binding transcription factor, Oct-1, is believed to regulate the expression of a number of ubiquitously expressed genes. These include genes which are expressed throughout the cell-cycle (eg. mRNA genes) and histone H2B genes, whose expression is tightly coupled to nuclear DNA synthesis at S-phase of the cell-cycle.

I have isolated and completely sequenced two *X. laevis* homologues of Oct-1. The high degree of relatedness of the two homologues indicates that these are likely to be copies of the same gene, which arose during the theoretical genome duplication event in *X. laevis* evolution. *X. laevis* and human Oct-1 display strong evolutionary conservation (85% identity over a stretch of 750 amino acids), which presumably means that *X. laevis* has a similar, if not identical function to human Oct-1.

Homology between human and *X. laevis* does, however, break down shortly before the N terminal end, at a point where alternate splicing is known to occur in human Oct-1 (W. Herr, pers. comm.). The full length *X. laevis* cDNA clones which I have isolated may represent a novel alternately spliced form of Oct-1.

Two octamer-binding proteins have been identified (in band shift assays) in *X. laevis* oocyte, embryo and tissue extract. Oct-1, and a second, previously unidentified octamer-binding protein which has been termed Oct-R, for octamer-related. Oct-1 does not bind to a degenerate octamer motif most often seen in *X. laevis* H2B promoters. Oct-R binds more strongly to this degenerate motif than the consensus motif, but only in the context of the H2B promoter, and does not bind either motif in another sequence context. This suggests that Oct-R may have a role in regulation of H2B transcription, although no direct evidence has been obtained.

Since Oct-1 is believed to stimulate the S-phase specific induction of histone H2B gene transcription the possibility that Oct-1 binding activity is cell-cycle regulated is of interest. *X. laevis* Oct-1 (and Oct-R) binding activity does not appear to be cell-cycle regulated.

Oct-1 and Oct-R are stored in the oocyte (partly in the cytoplasm), in an amount equivalent to at least 80 000 somatic cells. Histone protein and message are stored in the oocyte as part of the mechanism to provide enough histones to keep-up with the high rate of DNA synthesis in early *Xenopus* development. It is possible that histone gene transcription factors are stored for the same purpose.

By mutation of the octamer motif in the promoter of a *X. laevis* histone H2B gene promoter I have tentatively concluded that the octamer motif is required for the expression of a H2B gene (independently of DNA synthesis) in the oocyte. The H2B gene occurs in association with a H2A gene, as part of a divergently expressed gene pair. The octamer motif may be required for the expression of both H2B and H2A genes. The degenerate octamer motif contained in this H2B promoter does not bind efficiently to Oct-1 *in vitro*, but binds well to Oct-R, indirectly suggesting that Oct-R is required for the expression of the H2B gene.

A polyclonal antiserum raised against the N terminal domain of *X. laevis* Oct-1 reacts to proteins other than Oct-1 on Western blots of oocyte and embryo extract. These proteins, which are antigenically related to the N terminal domain of Oct-1, are entirely located in the cytoplasm of the oocyte, and entirely located in the nucleus of somatic cells. These proteins are synthesised during oogenesis, and stored in the oocyte in an amount equivalent to at least 100 000 somatic cells.

Abbreviations.

Ac	acetate.
Ad	adenovirus.
ATP, dATP,	adenosine triphosphate, deoxyadenosine
ddATP	triphosphate, dideoxyadenosine triphosphate.
bp	base pair.
BCIP	4-chloro-3-indoyl phosphate.
cDNA	complementary DNA.
Ci	Curie.
CIAP	calf intestinal alkaline phosphatase.
CTP, dCTP,	cytidine triphosphate, deoxycytidine
ddCTP	triphosphate, dideoxycytidine triphosphate.
DNA	deoxyribonucleic acid.
DNase	deoxyribonuclease.
DTT	dithiothreitol.
dTTP,	Thymidine triphosphate, deoxythymidine
ddTTP	triphosphate.
EDTA	Ethylene diamine tetra acetic acid.
GH	growth hormone.
GTP, dGTP,	guanosine triphosphate, deoxyguanosine
ddGTP	triphosphate, dideoxyguanosine triphosphate.
HACoCl ₃	hexamino cobalt trichloride.
HEPES	N-[2-Hydroxyethyl] piperazine -N'-[2-ethane sulphonic acid].
HSV	herpes simplex virus.

Ig	immunoglobulin
IgH	immunoglobulin heavy chain.
Igk	immunoglobulin kappa light chain.
IPTG	isopropyl-B-D-thiogalactosidase.
Kb	kilobase pairs.
KD	kilodaltons.
l	litre.
MES	2 [N-morpholino] ethane sulphonic acid.
ml	millilitre.
MOPS	3-(N-morpholino) propanesulfonic acid.
MS222	3-aminobenzoic acid ethyl ester.
ng	nanogram.
NBT	nitro-blue tetrazolium.
ORF	open reading frame.
pfu	plaque forming units.
Pg	picogram.
Pol	polymerase.
Prl	prolactin.
RNA	ribonucleic acid.
RNase	ribonuclease.
rpm	revolutions per minute.
rRNA	ribosomal RNA.
S	Svedberg.
SDS	sodium dodecyl sulfate.
snRNA	small nuclear RNA.
SV40	simian virus 40.

TCA	trichloroacetic acid.
Tris	Tris(hydroxymethyl)aminomethane.
tRNA	transfer RNA.
ug	microgram.
ul	microlitre.
UTP	uridine triphosphate.
v/v	volume/volume.
w/v	weight/volume.
X-gal	5-Bromo-4-chloro-3-indoyl-B-D-galactopyranoside.

Introduction

CHAPTER 1: INTRODUCTION.

1.1 The Octamer Paradox.

The octamer motif was first shown to be a transcriptional control element in the immunoglobulin kappa chain promoter (Falkner and Zachau (1984)). The octamer motif (ATTGCAAT) or its inverse complement was subsequently recognised as a conserved promoter/enhancer sequence element involved in the regulation of a range of genes with diverse patterns of expression. These genes include immunoglobulin heavy and light chain genes (the expression of which is restricted to B lymphocytes), U SnRNA genes (ubiquitously expressed) and histone H2B genes (ubiquitously expressed at S-phase of the cell-cycle). The octamer motif associated with these genes was shown to bind a common factor (Sive and Roeder (1986)). Following this there was an explosion of interest in how a the same motif, binding a common factor was a able to specify diverse patterns of gene expression. Prior to recognising the octamer motif as important the accepted model was that the DNA binding domain of a transcription factor conferred the specificity of transcriptional activation by attaching the activation domain to a particular promoter. This model cannot explain how a common promoter element is able to regulate diversely expressed genes. The 'octamer paradox' has been partially resolved by the identification of several factors which bind the octamer motifs associated with these genes.

Introduction

In the few years since their discovery, octamer-binding transcription factors have probably become the transcription factors about which most is known. Properties which have been investigated for the octamer-binding transcription factors, are applicable to other factors. These include regulation of diversely expressed genes (also seen with, for example, NF-kB which regulates immunoglobulin genes and a variety of genes in non-B cells; see Lenardo and Baltimore (1989)), flexible binding site requirements (also seen with, for example, NF-kB and CREB/ATF; see Lenardo and Baltimore (1989), Jones et al (1988)), and occurrence as a family of proteins binding to the same sites (also seen with, for example, fos/jun and CREB/ATF; see Ziff (1990)).

Following the cloning octamer-binding transcription factors it became apparent that that these proteins contain homeoboxes. The homeobox was identified as a conserved region between Drosophila homeotic proteins (Gehring (1987)). Mutations in homeotic genes transform certain parts, or an entire body segment into the corresponding structures of another body segment. Since homeobox genes have been cloned largely by classical genetic analysis or by homology to other homeobox genes, the target genes of these putative transcription factors are much less well defined than the targets of octamer-binding transcription factors. A study of octamer-binding transcription factors will no doubt provide an insight as to how homeobox genes

Introduction

in general regulate transcription and thereby specify segmental identity. It is becoming apparent that several Drosophila homeobox proteins bind to common sites (reviewed in Biggin and Tijan (1989)), and features discovered for the octamer-binding transcription factors, such as interaction with other proteins via specific residues in the homeobox, may be applicable to the mechanisms by which homeobox proteins in general differentially activate transcription through common binding sites.

This section describes the genes associated with the octamer motif, and their diverse patterns of gene expression for which the octamer motif is necessary. As already noted, the 'octamer paradox' has been partially resolved by the identification of several factors (using band shift assays) which bind to the octamer motifs associated with these genes. Here, I describe binding studies on the best characterised octamer-binding proteins (Oct-1 and Oct-2A/Oct-2B) using crude extracts and part-purified octamer binding proteins in relation to the expression of these genes. Studies on the differential abilities of Oct-1 and Oct-2 (including cloned Oct-1 and Oct-2) to activate transcription will be discussed in more detail in later sections, as will studies on other octamer binding proteins.

1.1.1 The octamer motif and immunoglobulin gene expression.

Transcription of immunoglobulin genes is both tissue

Introduction

specific (in B lymphocytes) and developmental stage specific (lower levels of expression in early B cell development). The transcription of immunoglobulin genes is controlled by both promoter and enhancer (downstream of the promoter, within introns) sequence elements, which contain a complex array of functionally active trans-acting factor binding sites (for recent review see Calame (1989)).

Octamer motifs occur within the heavy chain gene (IgH) promoter, the IgH enhancer, the kappa light chain gene (IgK) promoter, and the IgK enhancer (a non-consensus octamer) (Calame (1989), Currie and Roeder (1989)). An extensive series of transfection studies (carried out by numerous groups) using constructs containing Ig promoters/enhancers with normal and mutated octamer motifs have examined the role of the octamer motif in Ig transcription. The octamer motif in Ig promoters is absolutely required for gene expression (Eaton and Calame (1987), Falkner and Zachau (1984)) and deletion of the octamer motif within the IgH enhancer reduces gene expression (Gerster et al (1987), Lenardo et al (1987)). The octamer is also a determinant of the B cell specific expression of Ig genes. Six copies of the IgH enhancer octamer downstream of the β -globin gene are a B cell specific enhancer (Gerster et al (1987)) and synthetic octamer motifs inserted into the renin gene promoter (which is normally expressed only in highly specialized tissues) linked to the β -globin TATA box create a high activity

Introduction

lymphoid specific promoter (Dreyfus et al (1987)), as do octamer oligonucleotides upstream of the β -globin TATA box (Wirth et al (1987)).

A variety of proteins binding to the octamer motif have been identified using band shift assays. The best characterised octamer-binding proteins are Oct-1 (Singh et al (1986), Landolfi et al (1986), Gerster et al (1987), Fletcher et al (1987) - also known as OTF-1, OBP100, NPA1 and NFIII) and Oct-2A (Landolfi et al (1986), Gerster et al (1987), Staudt et al (1986), Scheidereit et al (1987) - also known as OTF-2, NFA2, IgNFA), which bind indistinguishably to the octamer motif. Oct-1 has been found in all cell types tested, and consequently implicated in the regulation of ubiquitously expressed octamer-containing genes. Oct-2A is predominantly expressed in lymphoid tissues (expression is also detected in the central nervous system, and other tissues), and is consequently implicated in the regulation of Ig genes. A second lymphoid specific octamer-binding protein has been identified (Oct-2B) (Gerster et al (1987), Schreiber et al (1988)) by band shift assays. Oct-2A/Oct-2B have subsequently been shown to be derived from the same gene by differential splicing (Hatzopoulos et al (1990)). Oct-2A itself may consist of three protein species (Scheidereit et al (1987)). Subsequent labelling of cells transfected with cloned Oct-2 with ³⁵S and ³²P has shown that different phosphorylated forms of Oct-2A exist (Tanaka and Herr

Introduction

(1990)), this may explain the multiple forms of Oct-2A identified by purification of the protein. The observation that bacterial lipopolysaccharide, which activates immunoglobulin expression in B cell lines, induces Oct-2 (but not Oct-1) binding activity (Staudt et al (1986)) adds weight to the theory that Oct-2 is involved in Ig gene transcription.

The binding patterns of extract from lymphoid and HeLa cells to octamer-containing DNA fragments originally described did not always consist simply of three bands (Oct-1, Oct-2A, Oct-2B) probably because the original band shift probes were restriction fragments with other binding sites (particularly Landolfi et al (1986), whose probe contained an heptamer motif, not then recognised as an octamer binding site - see later), and because some groups did not detect Oct-2B in their extracts (probably a technical problem).

The heptamer motif is a conserved sequence element adjacent to the octamer in the IgH promoter (Eaton and Calame (1987)), which contributes to the activity of the promoter. Both Oct-1 and Oct-2A bind to the heptamer motif, but only when the intact octamer motif is present and occupied (Landolfi et al (1988), Poellinger and Roeder (1989)). Binding studies with cloned Oct-2A have shown that the same region of Oct-2A is responsible for binding to the octamer and the heptamer (a highly degenerate octamer motif) (LeBowitz et al (1989)). Oct-2B also binds in a

Introduction

similar cooperative fashion (Kewler et al (1989)). This feature may have a role in the activation of IgH genes before Igk genes. Pre B cell lines have been shown to contain less Oct-2A (Staudt et al (1988)), and the cooperative binding may favour the activation of IgH in the presence of low levels of Oct-2A.

There is no simple relationship between Ig expression and Oct-2 expression. Oct-2 expression is detected (at lower levels) in cells which do not express Ig genes (see section 1.3.1). The role of Oct-2 in cells not expressing Ig genes is not clear. Also, IgH expression occurs in cell lines which do not express Oct-2 (Cockerill and Klinken (1990)). This serves to point to the involvement of other factors in the control of Ig gene expression.

Another aspect of Ig gene expression is extinction in hybrid cell fusions. Genes activated by the IgH enhancer are inactivated on the fusion of B cells with a T cell lymphoma cell line. Likewise, genes activated by a multimerised octamer (which is a B cell specific enhancer) are inactivated. This does not correspond to the loss of Oct-2 binding activity (Yu et al (1989)). Inactivation is dominant to activation by the SV40 enhancer, and so possibly not due to the loss of an IgH enhancer transactivator. This suggests that T cells contain a negative regulator, which associates with the octamer motif. There is other evidence of negative regulators acting via the octamer motif (Lenardo et al (1989)). An

Introduction

octamer-containing fragment from the IgH enhancer is active in non-lymphoid cells, but inactive in the F9 (embryonal carcinoma) cell line. However, mutation in the octamer renders the fragment an active enhancer. An F9 specific octamer binding activity (NFA-3) has been detected, and is a candidate for the negative regulator. However extinction of the Igk enhancer on fusion of a myeloma cell line with a fibroblast cell line is associated with decreased Oct-2 binding activity (and message) (Bergman *et al* (1990)). Similarly, B cells stably transfected with a construct containing the Igk promoter upstream of a β -globin reporter gene show extinction of the reporter gene on fusion with a fibroblast cell line (Junker *et al* (1990)). This extinction is dependent on the octamer motif (if replaced by NF1 or Sp1 binding sites there is no extinction). The extinction is overcome by transfection with cloned Oct-2A, and extinction is associated with the loss of Oct-2A/B message and protein.

1.1.2 The octamer motif and histone H2B gene expression.

The major class of histone genes in eukaryotes is replication dependent. Biosynthesis is closely linked to nuclear DNA synthesis at S-phase of the cell-cycle. A subclass of histone genes show a low level of constitutive expression. The regulation of the reiterated replication-dependent histone genes occurs at post transcriptional (control of mRNA stability, and regulation of 3' end

Introduction

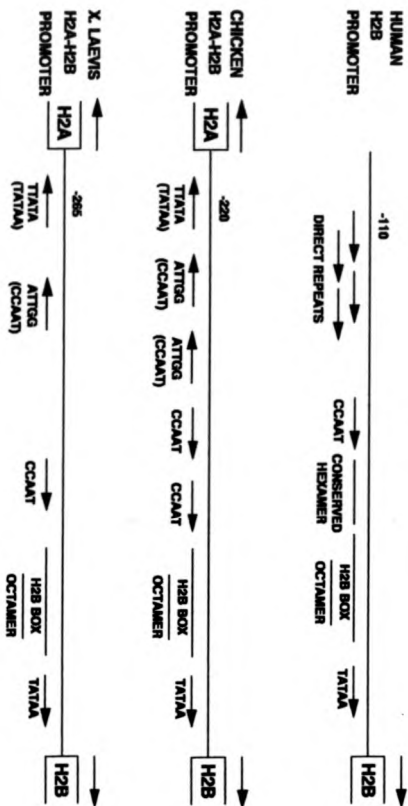
formation) and transcriptional levels (for review see Schumperli (1988)). The transcriptional regulation of histone subtypes in higher eukaryotes seems to depend on subtype-specific promoter elements. As I will describe in more detail below, H2B genes have an octamer motif in the promoter required for the S-phase stimulation of transcription. H1 genes have a subtype-specific consensus element, shown by mutational analysis to be required for the S-phase specific stimulation of H1 expression in transfected HeLa cells (Dalton and Wells (1988)). The level of a factor binding to this element is increased at S-phase of the cell-cycle (Dalton and Wells (1988A)). The arrangement of cis-acting sites in H4 genes is much more complex (see Dailey et al (1988) and references therein), but a subtype-specific consensus element has been identified and a binding factor purified. It is not clear from any mutational analysis that this element is required for the S-phase specific stimulation of transcription in vivo, but this factor is able to activate transcription from promoters containing the consensus element in vitro. No clear consensus element has been described for histone H3 promoters, but sequences required for the S-phase specific stimulation of transcription of a hamster H3 gene have been identified, and factors binding to these sequences show enhanced levels at S-phase of the cell-cycle (Artishevsky et al (1987)).

Another aspect of histone gene regulation is the

Figure 1.

The structure of histone H2B gene promoters.

Human, chicken and *X. laevis* histone H2B gene promoters are illustrated schematically (see LaBella *et al* (1988), Sturm *et al* (1988), Aldridge (1986)). In chicken and *X. laevis* the H2B gene generally occurs as a divergently expressed gene pair in association with a H2A gene. The boxes indicate the coding region for each gene, and the arrows indicate the direction of transcription. CCAAT boxes, TATAA boxes and H2B boxes (with core octamer) are illustrated, as are two sequence motifs (direct repeats and hexamer) specific to the human H2B promoter. Numbers refer to the approximate position relative to the H2B transcription start site.



Introduction

coordinate expression at S-phase of the cell-cycle. If subtype-specific elements are responsible for the S-phase specific stimulation of transcription, there must be a common mechanism for the expression/modification of binding factors (including Oct-1, if it regulates histone H2B genes). Histone genes are multicopy, and occur in clusters. In birds and mammals the clusters are 'randomly' organised, but in amphibians and lower eukaryotes the histone genes are generally organised into quintets of the 5 subtypes, and in some cases the quintets are strictly ordered (for review see Old and Woodland (1984)). The quintet organisation was once thought to have some regulatory significance (perhaps in coordinate control), but there is no evidence for this. Instead, the quintet organisation may have an evolutionary significance associated with the pressure to maintain equal numbers of histone gene subtypes on multiplication.

The structure of typical human, chicken and Xenopus laevis histone H2B gene promoters is illustrated in figure 1. In chicken and Xenopus laevis H2B genes generally occur associated with H2A genes as part of a divergently expressed gene pair. An octamer motif occurs in the H2B gene promoter, as the core of a longer consensus sequence known as the H2B box.

Transfection studies with reporter constructs containing a wild-type, and mutations of a, human histone H2B gene promoter have revealed the following facts (LaBella et al

Figure 2.

A sequence comparison of H2B boxes.

H2B boxes from X. laevis, chicken, mouse and human H2B gene promoters are aligned. A species consensus sequence is determined for X. laevis and chicken, since the sequence of several H2B promoters is available. Beneath the consensus sequence the number of times a particular base occurs is shown. Making the assumption that the single mouse and human sequences available are representative, an overall consensus is determined from these sequences and the chicken and X. laevis consensus sequences. Sequences are from LaBella et al (1988), Sturm et al (1988), Perry et al (1985), Moorman et al (1982), Wells (1986), Zhong et al (1983) and Aldridge (1986).

H2B BOX

<u>SPECIES</u>	<u>CLONE</u>	<u>OCTAMER</u>
X. laevis	X1HW23	CTGCCTC GTTTGCAT GGG
	X1HW11	CTGCCTC GTTTGCAT GGG
	X1H1	CTGCCTC GTTTGCAT GGG
	X1HW8	CAGCCTC GTTTGCAT GCG
	X1H3	GCCAGTG CTTTACAT GGG
	X1Hh1	CTGCCTT ATTTGCAT GGG
	X1HW7	CTACCTT ATTTGCAT GGG
	X1HW28	CTGCCTT ATTTGCAT GGG
	CONSENSUS	CTGCCTC GTTTGCAT GGG
	OUT OF 8	7667784 48887888 8788
Chicken	pCH3.3E	GCTTCTC ATTTGCAT AGAG
	pCH22.OB	TCTCCTA ATTTGCAT ACCG
	pCH3.5E	GCTGCTG CTTTGCAT AGCG
	pCH11.OEL	GCTGCTG CTTTGCAT AGCG
	pCH11.OER	GAACCTA ATCTGCAT ACAG
	CH1-10	CCTCTGT ATTTGCAT AACA
	CONSENSUS	GCTCCTN ATTTGCAT AGCG
	OUT OF 6	445355- 46566666 6345
Human	pHh4c	TCACCTT ATTTGCAT AAGC
Mouse	pMH2B	GCTTGAC GTTTGCAG ACTC
OVERALL CONSENSUS		GCTCCTC ^A TTTGCAT AGCG
OUT OF 4		2323332 24444443 3222

Introduction

(1988)). The direct repeats, hexamer and CCAAT box can all be deleted without affecting the stimulation of transcription on entry into S-phase of the cell-cycle (which is normally approximately 5 fold). Mutation of the octamer, however, completely abolishes this stimulation. The direct repeats and CCAAT box are required for both basal and S-phase levels of expression. Mutation/deletion of these elements gives a lower basal activity, which is stimulated approximately 5-fold at S-phase, giving a lower S-phase level of activity. The octamer is not required for basal expression. The conserved hexamer (which makes a small contribution to expression in in vitro assays (see Sive et al (1986)) has no effect on expression in these transfection assays.

The octamer motif has been shown to be required for the steady state levels of chicken H2B and H2A transcription in transfected cells (Sturm et al (1988)). Here the genes occur as a divergently expressed pair, and octamer mutations decrease the expression of both genes. Separation of the two genes reduces expression of both genes more than the octamer mutation, and deletion of the CCAAT box most proximal to the H2B gene decreases expression of both genes. Consequently the genes may have other promoter elements in common.

As already noted, in H2B gene promoters the octamer occurs as the core of a longer consensus sequence, known as the H2B box (see figure 2, sequences from LaBella et al (1988),

Introduction

Sturm et al (1988), Perry et al (1985), Moorman et al (1982), Wells (1986), Zhong et al (1983), Aldridge (1986)). The H2B box is always referred to as being conserved between species. The octamer (ATTGTCAT) is well conserved between H2B promoters within a particular species and between species, although degenerate octamer motifs occur. The sequence GTTTCAT is the most common degenerate octamer motif, and GTTTCAT is the most common octamer motif in Xenopus laevis histone H2B gene promoters. However, conservation of sequences flanking the octamer is most striking between genes within a particular species. This is particularly true for Xenopus laevis, where strong conservation is seen over a stretch of 19 bases. Conservation of sequences flanking the octamer between species is, however, not extensive. In figure 2, X. laevis (from 8 sequences) and chicken (from 6 sequences) H2B box consensus sequences have been determined. These sequences are not well matched outside the octamer, and do not match well to the human and mouse H2B boxes outside the octamer. Only one human and one mouse H2B box sequence is available, and making the assumption that these sequences are representative, an overall consensus from these sequences and the chicken and X. laevis consensus sequences is shown in figure 2.

Human (Sive and Roeder (1986)), hamster (Ito et al (1989)), and chicken (Dalton and Wells (1988A)) Oct-1 have been shown to bind to the H2B octamer, and I have shown

Introduction

that Xenopus laevis Oct-1 binds to the H2B octamer. A simple model for the S-phase specific stimulation of H2B transcription is that the ubiquitous octamer binding factor (Oct-1) is the transactivator responsible for activation, and that levels of binding activity are enhanced at S-phase of the cell-cycle. The data concerning the cell-cycle regulation of Oct-1 binding activity are contradictory (and will be discussed in more detail later). Briefly, there is no evidence that in human cell lines Oct-1 binding activity is cell-cycle regulated. In avian cells Oct-1 binding activity does not vary through the cell-cycle (Dalton et al (1988A)), whereas in the same cells the factor binding to the H1 subtype-specific consensus element displays elevated levels at S-phase of the cell-cycle. Perhaps this is to be expected since Oct-1 is also implicated in the control of ubiquitously expressed genes which are not cell-cycle regulated (eg. snRNA). Cell-cycle regulated modification of Oct-1, or interaction with other factors could account for S-phase specific stimulation of H2B genes by Oct-1 (discussed in more detail later). However, in hamster cells, the level of Oct-1 binding activity is elevated at S-phase of the cell-cycle (Ito et al (1989)).

1.1.3 The octamer motif and snRNA gene expression.

The transcriptional control elements of U snRNA genes are distinct from other RNA polymerase II transcribed genes (for review see Parry et al (1989)). The genes have no TATA

Introduction

box, but a PSE (proximal sequence element) which serves a similar purpose, being necessary for transcription and specifying the transcriptional initiation site. Upstream of the PSE is an enhancer element, the DSE (distal sequence element). The DSE contains octamer motif(s) in association with auxiliary binding sites. Deletion of the octamer-containing DSE eliminates snRNA expression, which can be partially restored by replacement of the DSE with an octamer motif (Parry et al (1989), Mattaj et al (1985)). Related experiments will be discussed in much more detail later. Since the octamer motif is important for U snRNA expression, and the genes are generally ubiquitously expressed, Oct-1 is thought to be the transcriptional regulator.

The U2 snRNA octamer is the consensus octamer (ATTGCGAT), which has been shown to bind Oct-1 and Oct-2. The human, rat and chicken U1 snRNA genes have a degenerate (but conserved) octamer motif (ATCTACAT) (Kemp and Latchman (1988)). The same is true for the U3 snRNA octamer, although the octamer is again distinct (ATTAGCAT). Most snRNA genes are believed to be ubiquitously expressed. However, developmentally regulated U1 and U4 snRNA genes occur in Xenopus, and the U1 genes have distinct octamer motifs (Krol et al (1985), Lund and Dahlberg (1987)).

U6 snRNA genes are transcribed by RNA polymerase II and III. Downstream of the DSE and PSE the U6 promoter contains a sequence resembling the pol II TATA box. This sequence

Introduction

element is required for pol III transcription (Parry et al (1989)). Pol III transcripts are functional, but pol II transcripts are of uncertain significance. Co-injection of the U2 DSE with the U6 gene into Xenopus oocytes diminishes U6 transcription (Carbon et al (1987)). This competition effect can be relieved by mutations in the U2 DSE octamer motif. Consequently, U2 and U6 DSE's seem to bind a common oocyte octamer factor required for the pol III-dependent transcription of U6 anRNA genes.

1.1.4 The SV40 enhancer contains an octamer motif.

Mutational analysis of the simian virus 40 (SV40) enhancer has shown that it is composed of multiple sequence elements which act synergistically to generate enhancer activity (for review see Jones et al (1988)). The sph motif is one such sequence element. Two direct repeat copies of the sph motif occur, and at the junction of these copies an octamer motif is present. The octamer motif is a 7/8 match to the consensus (CTTTGCAT, not ATTTGCAT). The sph motif is required for enhancer activity in HeLa cells, however the octamer is not. The reverse situation is true for B cells (Davidson et al (1986)). Correspondingly, interaction with the octamer motif (and not the sph motifs) can be detected by DNase footprinting with B cell extract, the reverse situation being true with HeLa cell extract. This leads to the model that Oct-2 binds to and activates the enhancer in B cells, whereas sph binding factors perform this role in

Introduction

HeLa cells. Such models will be discussed in more detail later. The situation is complicated by band shift assays. Band shift assays detect factors binding to the octamer motif in both B cells and HeLa cells, but do not detect *aph* binding factors (Rosales *et al* (1987)). Rosales *et al* (1987) detect 4 octamer binding proteins interacting with the SV40 enhancer octamer, and the IgH enhancer octamer. The previously reported B cell specific Oct-2A (here called Oct-B2) is identified. Oct-2B is not. It is suggested that the ubiquitous binding activity, Oct-1, is in fact 2 distinct proteins of similar mobility - one being B cell specific, and the other absent from B cells. This is on the basis of apparently different ability to protect the IgH and SV40 enhancers in footprinting assays, and since elution from an ion exchange chromatography column occurred at different salt concentrations. These differences have not been reported by other workers, and perhaps could represent differences in Oct-1 modification. Studies with cloned Oct-1 indicate that the message is ubiquitously expressed (Sturm *et al* (1988A)). Rosales *et al* (1987) also detect a minor ubiquitous binding activity, which has not been reported elsewhere and could represent a degradation product. This reflects the general complexity and confusion when comparing the band shift data from from the many groups working in this area, particularly since a second adjacent octamer site has been identified in the SV40 enhancer (a situation reminiscent of the IgH heptamer).

Introduction

HeLa cell Oct-1 binds both sites (GTTTGCAT and ATTGAGAT) (Sturm et al (1987), Baumruker et al (1988)) giving an upper band in band shift assays corresponding to both sites being occupied. This phenomenon was not reported by Rosales et al (1987).

1.1.5 The octamer motif and the expression of HSV alpha (immediate early) genes.

Herpes simplex virus (HSV) alpha genes contain a conserved TAATGARAT motif (R is a purine) in the promoter. Functional studies have identified this motif to be necessary for the induction of alpha genes following infection (Gellman and Silverstein (1987), Gaffney et al (1985), Macken and Roizman (1982)) In most cases an octamer motif is found overlapping this motif. An example of such an element (from the HSV-1 ICPO gene) is shown below.

TAATGARAT
7/8 octamer
ATGCTAATGATAT
TACGATTACTATA

A viral protein VP16 (also known as Vmw65, alpha IIF) present in the viral particle is necessary for the induction of alpha genes via the TAATGARAT motif (Campbell et al (1984)).

However, binding of VP16 to DNA cannot be demonstrated (Marsden et al (1987)). In band shift assays with the sequence above as probe and extract from virally infected HeLa cells, an additional complex, of lower mobility to the

Introduction

Oct-1 complex, is formed (Preston et al (1988), O'Hare and Goding (1988), McKnight et al (1987), Gerster and Roeder (1988)). Oct-1 and VP16 bind together to the alpha gene extended octamer motif, VP16 binding only when Oct-1 is present. Oct-1 can bind alone. Oct-1 and VP16 together (but not Oct-1 alone) seem to activate reporter genes containing an alpha gene consensus element in the promoter, in HeLa cells (O'Hare et al (1988)). This will be discussed in more detail later. Oct-1 is also able to bind (albeit with lower affinity) to the TAATGARAT motif without associated octamer, which is the arrangement seen in some alpha gene promoters (apRhys et al (1989)).

1.1.6 The octamer motif and the expression of other genes.

The 7SK RNA gene has an unusual promoter resembling the the U6 snRNA promoter (Murphy et al (1987), Murphy et al (1989)). There is a TATA box (normally a requirement of RNA pol II transcription) just upstream of the transcription start site, upstream of this an element resembling the PSE of U snRNA genes, and further upstream at least three degenerate octamer motifs. The 7SK RNA gene is transcribed by RNA polymerase III. Mutational analysis of the promoter indicates that the octamer motifs operate cooperatively to activate 7SK transcription in vitro. The consensus octamer motif is able to substitute for the degenerate upstream octamer motifs. Purified Oct-1 and Oct-2 activate transcription from this promoter in vitro, and are able to

Introduction

bind to the very degenerate octamer motifs associated with this gene.

An octamer motif occurs in the promoter of the human γ -globin gene (Gumucio et al (1988)). Expression of γ -globin in erythroid cells is normally repressed at birth. One mutation associated with elevated γ -globin expression in the adult is in the octamer motif of the γ -globin promoter. Oct-1 and Oct-2 like binding activities from an erythroleukaemia cell line bind to the γ -globin octamer. Binding is inhibited by the mutation associated with elevated γ -globin expression. Here, it seems that octamer binding proteins may act as repressors of transcription. Erythroid-specific factors bind to activator sites found in close association with the octamer. Another site (significance to expression untested) overlapping the octamer binds an erythroid-specific factor. Octamer factor binding may inhibit the association of activator proteins with the γ -globin promoter.

Octamer motifs occur upstream of other genes, and many more octamer-associated genes will no doubt be identified, particularly in view of the fact that multiple octamer binding proteins have now been identified (discussed later) and because Oct factors are able to bind remarkably degenerate octamer motifs (the flexible binding requirements of Oct factors will be discussed later), for example the IgH heptamer, in the 7SK promoter and the HSV alpha gene TAATGARAT motif. This point can be illustrated

Introduction

by the finding that a repressor site in the human c-myc promoter (expression of c-myc is associated with the growth state of cells) forms two complexes in band shift assays (Takimoto et al (1990) and references therein). These complexes were identified by competition with known binding sites. The AP1 (fos/jun complex, a transactivator of viral and cellular genes) binding site competed out one complex, and the consensus octamer motif the other complex. AP1 and Oct-1 protect highly overlapping regions in the repressor element, and these represent very degenerate binding sites for the two factors, which would not have been readily identified by sequence analysis.

1.1.7 The octamer motif and adenovirus replication.

Octamer motifs are conserved in the adenovirus (Ad) origin of replication (Pruijin et al (1987), Verrijzer et al (1990) and references therein). The Ad2 origin contains the octamer ATTATCAT, and the Ad4 origin the consensus octamer, ATTGCGAT. Both these octamer motifs bind Oct-1 (the factor binding was originally called NFIII, but is identical to Oct-1), and purified Oct-1 is able to stimulate DNA replication from the Ad origin in vitro. It is interesting to note that a common factor (Oct-1) is implicated in the activation of DNA replication and the activation of histone H2B gene transcription, which is tightly coupled to DNA synthesis.

Introduction

1.2 Models of transcriptional activation by Oct-1/Oct-2 from transfection studies in B and non-lymphoid cells.

The evidence that Oct factors were involved in gene regulation was originally circumstantial. Octamer elements (as already described) were demonstrated to be required for the expression of diversely expressed genes, presumably by association with trans-acting factors. Oct-1 and Oct-2 were discovered, and found to bind indistinguishably to the octamer motifs within the promoters/enhancers of these genes. Oct-2 was originally believed to be restricted to lymphoid tissue (Oct-2 has subsequently been found at lower levels in other tissues) was said to be the octamer-binding transcription factor regulating Ig genes, and Oct-1 (because of its ubiquitous distribution) was the octamer-binding transcription factor regulating ubiquitous genes. The fact that ubiquitously expressed genes showed different patterns of expression (H2B genes are expressed at S-phase of the cell-cycle) was explained by hypothetical modification of Oct-1, or association with another factor(s).

This section will describe the promoter/enhancer constructs that have been used for transfection studies in B cells (expressing Oct-1 and Oct-2) and non-lymphoid cells (expressing Oct-1) to analyse the mechanism of transcriptional activation by Oct-1 and Oct-2. It should be borne in mind that the models derived from these experiments make the assumption that Oct-1 and 2 are the only octamer-binding transcription factors present, and are genuine in

Introduction

vivo transactivators. Studies on transcriptional activation using cloned Oct-1 and Oct-2 will be discussed in a later section.

1.2.1 Oct-1 and Oct-2 are transcriptional activators.

An important demonstration, and key to any model of transcriptional activation by Oct-1 and Oct-2, was that purified Oct-1 and Oct-2 were able to activate transcription.

As described earlier, the octamer motif is required for the S-phase specific stimulation of histone H2B transcription in vivo. In vitro transcription assays using the human H2B promoter show that octamer mutations have no effect on transcription using G2-phase HeLa cell extract (Fletcher et al (1987)). However, these mutations diminish the transcription observed with S-phase extract. So, in line with the in vivo transfection data, there seems to be an enhancement of an octamer-dependent transcription activation activity at S-phase of the cell-cycle. However, it is not clear whether or not HeLa cell Oct-1 binding activity is enhanced at S-phase of the cell-cycle. In other systems the data concerning the levels of Oct-1 through the cell-cycle are contradictory (see section 1.1.2). In an avian cell-line Oct-1 binding activity is enhanced at S-phase, whereas in a hamster cell line the level of Oct-1 binding activity is constant through the cell-cycle. Ion exchange and DNA affinity chromatography have been used to purify HeLa cell Oct-1. The purified, approximately 90 kd protein is able to

Introduction

stimulate in vitro transcription from the intact H2B promoter, but not from a promoter with mutations in the octamer motif. The single 90 kd protein band cut from an SDS gel after electrophoresis of purified Oct-1 is an active octamer binding protein. The band shift seen is identical to that seen with crude extract.

B cell Oct-2A has also been purified by ion exchange and DNA affinity chromatography (Scheiderait et al (1987)). Octamer mutations in the kappa promoter diminish activity in B cells (but not in HeLa cells). Octamer-factor-depleted extract (by oligo affinity, and showing no Oct binding activity in band shifts) shows no difference in the level of transcription from Oct+ and Oct- kappa promoters. However, adding purified Oct-2A stimulates transcription from Oct+ (and not Oct-) promoters. SDS-PAGE of affinity purified Oct-2A shows that it consists of three major bands of similar size (62, 61, 58.5 kd). Each of these proteins isolated from a gel is active in octamer binding, as assayed by band shift assays. Differentially phosphorylated forms of cloned Oct-2A have subsequently been shown to have different electrophoretic mobility on SDS-PAGE (Tanaka and Herr (1990)). This is potentially the source of the three Oct-2A species observed here.

A question which arises is 'Why doesn't Oct-1 stimulate Ig transcription in non-B cells?' As will be described in the following section, Oct-1 may have a weaker activation domain than Oct-2, which requires other factors to activate the TATA

Introduction

box transcription complex. However, the answer may be, in part, quantitative (LeBowitz et al (1988)). Low levels of octamer-dependent transcription are observed in transfected non-B cells, and similarly octamer-dependent Ig transcription occurs at low levels in vitro, with extract from non-B cells. Also affinity purified Oct-1 added to HeLa cell extract proportionally stimulates octamer-dependent transcription from an Ig promoter (it is also possible that an inhibitor is titrated out by the Oct-1 added).

1.2.2 Transfection of SV40 enhancer-containing constructs into B and non-lymphoid cells.

As already noted the octamer motif linked to a heterologous promoter forms a B cell specific promoter/enhancer (see section 1.1.1), leading to the suggestion that Oct-2, and not Oct-1, is able to activate such promoters, and that Oct-2 is a determinant of the B cell specific expression of Ig genes. Such experiments were extended further by a series of constructs containing the SV40 enhancer Sph motifs linked to heterologous promoters (Tanaka et al (1988)).

The SV40 enhancer (section 1.1.4) contains two direct repeat copies of the Sph motif. At the junction of these repeats an octamer motif is present. Three mutations were made in the Sph region to: (1) inactivate the Sph motif, which is required for HeLa cell activity of the enhancer, but does not inactivate the octamer motif, which is required for B cell activity of the enhancer [sph⁻oct⁺]; (2) inactivate the

Introduction

octamer motif, giving less activity in B cells (and less Oct factor binding), but not the Sph motif (activity in HeLa cells unaffected) [sph⁺oct⁻]; (3) inactivate both motifs, giving reduced activity in both cell types [sph⁻oct⁻]. Six copies of the sph motifs were cloned upstream of the β -globin TATA box, 2.2 kb downstream of the β -globin start site, and upstream of the U2 snRNA PSE, with the octamer-containing DSE removed (see section 1.1.3). Transcription of these constructs was assayed after transfection into HeLa and NS1 B cells.

In line with previous results, sph⁺oct⁺ and sph⁻oct⁺ are active in B cells, upstream or downstream of the β -globin start site (although 12 times less efficiently downstream). sph⁺oct⁻ and sph⁻oct⁻ are inactive in B cells. sph⁺oct⁺ and sph⁺oct⁻ are active when associated with the β -globin promoter in HeLa cells. sph⁻oct⁺ and sph⁻oct⁻ are inactive. Sph and Oct motifs are, therefore, separable and overlapping motifs, the sph motif being required for HeLa cell activity, and the octamer motif for B cell activity.

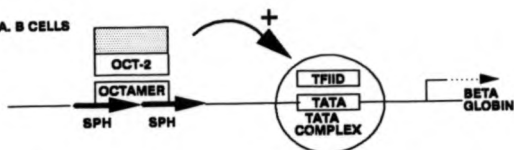
In these assays the U2 genes give two transcripts. A transcript from the correct start site which terminates at the conserved 3' box (which directs correct 3', unpolyadenylated end formation). There are also incorrect transcripts from cryptic promoters, which terminate at an adenovirus polyadenylation site found in the vector downstream of the 3' box. These are more like mRNA transcripts, and the effects of the enhancer mutations on the

Figure 3.

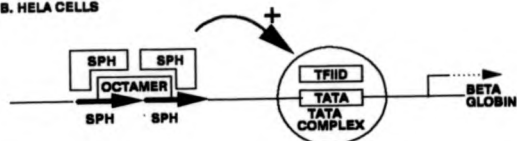
A model for transcriptional activation by Oct-1 and Oct-2.

SV40 enhancer sph motifs inserted upstream of the beta-globin TATA box or U2 snRNA PSE, as described in the text, are illustrated schematically (see Tanaka *et al* (1988)). Six copies were inserted, but for simplicity only a single copy is shown. In all cases the octamer motif at the junction of the sph motif is intact, and in all cases except (C), where the motifs are mutated, the sph motifs are intact. The constructs were transfected into the cell types indicated, and in the case of (E) co-transfected with a VP16 expressing construct. The boxes over the sph/oct motif represent the association of Oct-1, Oct-2 and sph binding factors. Oct-2 may have a stronger activation domain, illustrated by a larger stippled box. In (E) VP16 associates with the Oct:DNA complex, providing an acidic activation domain (stippled box). TATA box and PSE transcription initiation complexes are illustrated by a circle or square, respectively. The TATA box binding protein, TFIID is illustrated. The ability of the sph/oct complex to activate the transcription initiation complex is indicated by + or -. An arrow indicates that transcription occurs, and a cross that transcription does not.

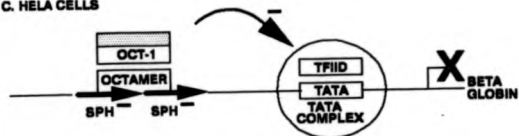
A. B CELLS



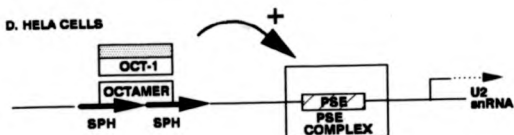
B. HELA CELLS



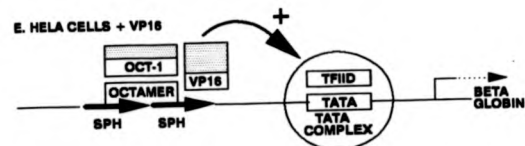
C. HELA CELLS



D. HELA CELLS



E. HELA CELLS + VP16



Introduction

level of the incorrect snRNA transcripts is the same as observed for the β -globin transcripts. However, the promoter mutations have a different effect on the level of the correctly initiated snRNA transcripts in HeLa cells. In HeLa cells correct transcripts are activated by sph^+Oct^+ and sph^-Oct^+ , whereas sph^+Oct^- and sph^-Oct^- are inactive. Upstream of the β -globin promoter, in HeLa cells, sph^-Oct^+ is inactive and sph^+Oct^- is active. So, in HeLa cells the octamer motif is able to activate the U2 promoter, but not the β -globin promoter.

Making the assumption that Oct-1 is responsible for transcriptional activation from octamer elements in HeLa cells, and that Oct-1 and Oct-2 are responsible for transcriptional activation in B cells the model of transcriptional activation illustrated in figure 3a-d is derived from this data. Oct-2 binds to the SV40 octamer in B cells and activates transcription from the TATA box, whereas in HeLa cells sph factors bind to the SV40 enhancer to activate transcription from the TATA box. When the sph motif is inactivated, Oct-1 binding in HeLa cells is unable to activate transcription from the TATA box. However, Oct-1 binding to the SV40 enhancer is able to activate transcription from the PSE in HeLa cells. Sph factor binding is unable to activate transcription from the PSE.

In HSV alpha genes a TAATGARAT motif is responsible for activation of transcription via the viral transactivator VP16 (see section 1.1.5). This motif frequently occurs in

Introduction

conjunction with an overlapping octamer, and is itself a degenerate octamer motif able to bind Oct-1. VP16 itself does not bind to DNA, but associates with Oct-1:DNA complex. In HeLa cells, in the presence of VP16 the normally inactive $\text{sp}^{\text{H}}\text{Oct}^+$ enhancer activates β -globin transcription. However, only in the proximal position. This is illustrated in figure 3e. A simple model, outlined in figure 3, is that Oct-2 presents a strong activation domain able to activate the TATA box transcription initiation complex. Oct-1 has a weaker activation domain, and is only able to activate the TATA complex in conjunction with the acidic activation domain of VP16. However, the Oct-1 activation domain is sufficient to activate transcription from the PSE complex. Recently the VP16 transactivator protein has been shown to interact selectively and strongly with TFIID (a TATA box binding protein and component of the TATA box transcription initiation complex) via its acidic activation domain (Stringer et al (1990)).

O'Hare et al (1988) have shown that an intact HSV alpha gene consensus motif with overlapping octamer linked to a reporter gene can be activated in HeLa cells by co-transfection with VP16. They report that both Oct-1:VP16:DNA complex formation and activation by VP16 requires the GARAT part of the motif (which flanks the octamer). However, VP16 can activate transcription from the SV40 octamer, the H2B octamer and the U3 snRNA octamer (see later) suggesting that the GARAT portion of the motif seen in some HSV alpha gene promoters is

Introduction

not a strict requirement, and VP16 may have flexible recognition requirements.

1.2.3 VP16 and the activation of U snRNA genes.

The U3 snRNA DSE octamer is a non-consensus octamer, similar to the HSV alpha gene consensus element which is required for transactivation by VP16 (Kemp and Latchman (1988)). The U1 snRNA DSE octamer is also a non-consensus octamer, but dissimilar from the alpha gene consensus. Co-transfection with VP16 stimulates U3 but not U1 transcription. U3 induction by VP16 requires the octamer motif, and the U3 octamer added to the U1 promoter (with its own octamer-containing DSE removed) renders it inducible by VP16. Since removal of the DSE reduces U1 expression, and this is partially restored by adding back either the U1 or U3 octamer it seems that the U1 and U3 elements must bind distinct octamer factors, one of which can be activated by VP16. In line with this argument is the observation that the U1 octamer, unlike the perfect octamer, linked to the β -globin promoter does not create a B cell specific promoter (Wirth et al (1987)).

1.2.4 VP16 and the activation of H2B transcription.

Transfection of cells with a VP16 encoding plasmid stimulates endogenous histone H2B transcription (Latchman et al (1989)). However, infection with HSV (which introduces VP16) represses H2B transcription. This suggests that HSV

Introduction

genes and H2B genes compete for VP16 by a common element. H2B gene promoters contain an octamer motif, and HSV alpha gene promoters a TAATGARAT motif (which itself binds Oct-1) which sometimes overlaps an octamer motif. Since VP16 is known to interact with the Oct-1:DNA complex, the suggestion is that during infection the alpha gene consensus competes for Oct-1 repressing H2B transcription, whereas when only VP16 is present it interacts with the Oct-1 to stimulate H2B transcription from the octamer motif. In line with this, the HSV alpha gene consensus and the U3 snRNA octamer motif co-transfected with VP16 overcome H2B stimulation by VP16. In contrast, co-transfected U1 octamer motif does not overcome induction of H2B transcription by VP16. This is further evidence that U1 and U3 octamer motifs bind distinct factors (see section 1.2.3).

Transfection of the H2B octamer motif alone (without VP16) represses H2B transcription, whereas transfection with the HSV alpha gene consensus does not. This suggests that VP16 promotes interaction of the common H2B promoter - alpha gene consensus octamer binding factor (presumably Oct-1) with the HSV alpha gene consensus.

VP16 lacking its acidic activation domain is unable to stimulate H2B transcription, but in the presence of the co-transfected alpha gene consensus is able to repress transcription. This implies that different domains of VP16 are responsible for the activation of transcription and interaction with the octamer-binding factor.

Introduction

A summary of the relevant points regarding the involvement of the octamer motif and Oct-1 in H2B transcription is: (1) the human H2B promoter octamer motif (the core of a longer consensus sequence known as the H2B box) is required for the S-phase stimulation of H2B transcription in vivo, and this stimulation is independent of other promoter elements; (2) H2B genes are ubiquitously expressed, and since Oct-1 is ubiquitous it is a good candidate for the octamer-binding transcription factor involved; (3) extract from S-phase HeLa cell nuclei, and not G2-phase, is able to stimulate octamer-dependent transcription of H2B genes in vitro; (4) affinity purified HeLa cell Oct-1 is able to stimulate octamer-dependent H2B transcription in vitro; (5) there is no evidence that Oct-1 binding activity is cell-cycle regulated in HeLa cells (human), in other species the evidence is contradictory; (6) octamer mutations in the human H2B promoter with other control elements inactivated (essentially an oct-TATA promoter) decrease expression in transfected HeLa cells, whereas mutations in the SV40 octamer upstream of the β -globin TATA box do not decrease expression in transfected HeLa cells; (7) VP16 is able to activate H2B transcription in vivo, in an octamer-dependent manner. VP16 interacts with the Oct-1:DNA complex, and activates the SV40 octamer - β -globin construct in HeLa cells.

This is compelling evidence for the involvement of Oct-1 in H2B transcription. However, the problem that Oct-1 is also involved in the transcription of ubiquitously expressed genes

Introduction

which are not cell-cycle regulated remains. One possible explanation is that Oct-1 undergoes cell-cycle regulated modification. Oct-1 is known to be phosphorylated and O-glycosylated (Tanaka and Herr (1990), Murphy et al (1989)). Another explanation is that Oct-1 undergoes periodic interaction with another factor, which is itself cell-cycle regulated. In favour of this explanation is the finding that the SV40 enhancer octamer upstream of the β -globin TATA box is inactive, whereas the octamer in the H2B promoter with other promoter elements inactivated (essentially an octamer - TATA promoter) is active in HeLa cells. This suggests that the sequence context of the octamer in the H2B promoter (and not in the SV40 enhancer) allows activation by Oct-1. This could be the spatial arrangement of the two promoters (which is different), but could also be the flanking sequences of the octamer in the conserved H2B box. The flanking sequences could facilitate interaction of Oct-1 with another cell-cycle regulated factor. A cellular analogue of VP16 could be such a factor in view of the findings that VP16 activates the SV40 octamer - TATA promoter and stimulates cellular H2B transcription in an octamer-dependent manner.

1.3 Studies on cloned Oct-1 and Oct-2A/Oct-2B.

The cDNAs encoding Oct-1 and Oct-2A/Oct-2B have been cloned and characterised. This section describes the cloning of these factors, information which has arisen from examining their sequence, and work which has been done to determine the

Introduction

regions of the proteins responsible for functions such as DNA binding, transcriptional activation and interaction with other proteins.

1.3.1 Cloning Oct-1 and Oct-2A/Oct-2B.

Human Oct-1 was cloned by South-Western blotting (screening a cDNA expression library with radiolabelled DNA binding site) (Sturm et al (1988A)), and the cDNA was identified to encode Oct-1 by the binding specificity of the product of the clone, by the fact that an antibody raised to the clone reacts with Oct-1 in band shift assays and by the ubiquitous distribution of the message from which the cDNA was derived. The distribution of the message was determined by RNase protection assays, and the probe used detected an additional smaller band with B cell RNA indicating a possible sequence homology between Oct-1 and Oct-2.

Human Oct-2A was cloned by South-Western blotting and by screening with oligonucleotides, the sequence of which was derived from the protein sequence of purified Oct-2A (Staudt et al (1988), Clerc et al (1988), Ko et al (1988), Scheidereit et al (1988), Muller et al (1988)). The cDNA was identified to encode Oct-2A on the basis of the distribution of the message from which it is derived (Oct-2A cDNA used as a probe for Northern blots detected transcripts in RNA from B cells, but not from non-lymphoid cell lines), binding specificity of the product of the cDNA and mobility of the product on SDS-PAGE. The transcripts detected on a Northern

Introduction

blot were of the sizes 6, 3.9, 3.3 and 2 kb, suggesting the potential for related Oct-2 messages, perhaps generated by alternate splicing.

Mouse Oct-2A and Oct-2B were cloned by homology to the human Oct-2A clone (Hatzopoulos et al (1990)). Mouse and human Oct-2A differ by only a few amino acid substitutions. The Oct-2B clone differs from Oct-2A as follows: a 48bp deletion (16 amino acids), and an insertion of 74 bases close to the 3' end which opens up a new reading frame, so that the last 12 amino acids of Oct-2A are replaced by 132 new residues. The product of the Oct-2B clone gives bands of the same mobility as Oct-2B in band shifts, and the same mobility as purified Oct-2B on SDS-PAGE. The mouse Oct-2 genomic clone indicates that Oct-2A and Oct-2B are alternately spliced products of the same gene.

RNase protection assays, using a probe which can distinguish between Oct-2A and Oct-2B transcripts, have been used to examine the distribution of the messages encoding the 2 factors. Oct-2A/B is present at much higher levels in lymphoid tissue, but also occurs in kidney, testis (an extra, smaller protected fragment is also detected indicating an additional protected transcript), intestine, brain, F9 embryonal carcinoma cells and embryos. The relative amounts of Oct-2A and Oct-2B vary. In situ hybridisation indicates that Oct-2 expression is widespread in the developing nervous system, with regions of higher expression. In the adult brain expression is restricted to distinct areas.

Figure 4.

The POU domain.

- (A) A sequence comparison of the POU domains of Oct-1 and Oct-2 (human), Pit-1 (rat) and unc-86 (*C. elegans*). Sequences are aligned to Oct-1. - represents a residue conserved with Oct-1, and * a gap made in the sequence for maximum alignment. The location of the POU specific and homeo box sub-domains is shown, as are the location of sub-regions within these sub-domains. The A, B, and WFC sub-regions are particularly well conserved. (See Herr et al (1988)).
- (B) The location of the POU domain within the proteins. The numbers refer to the number of amino acids in the protein
- (C) An alignment of the Oct-1, Oct-2 and antennapedia homeo boxes. Vertical lines represent a conserved amino acid. The location of possible alpha helices is shown. By analogy to bacterial repressor proteins, I, II, A and B are possible contact sites with DNA. (See Garcia-Blanco et al (1989)).

A

POU-SPECIFIC BOX

Oct-1 EEPDLE****ELRQFAKFKQRRIKLCFQGDVGLANGCLYGV**NDPSQTTISNPEALNTLSFKNMCKLPLLEKMLDAENLSS
 Oct-2 -----R-----
 Pit-1 --I-MDSPAIR-----NE-V-----Y--TN-E-LAAVH--**SE-----C--N-Q-----A--AI-S--EE-QYGA
 unc-86 PPT-MDT*DPKQ--T-EH-----V--A--K-LAH-KMPQVCSL--S--C--S-T--HN--VA--I-HS--EK-E**A

A

B

POU HOME BOX

Oct-1 DSSLSSPSPALNSP**CIFGL*SSRRKKRTSIEENIVALEKSPLENOKPTSEETITNADLNKEEVIIRVUCNRQKQEKRIKPPSS
 Oct-2 -----P--NO-S--SL--PD-----V--F-----A-----LI--R--H--L-----V-----R--VKTSLN
 Pit-1 LYNEK*****VGAN+E-K-R--T-SIAKD--RH-G-HS--S-Q--NM--EE--L-----Q--Q--Q--DPRSQF
 unc-86 MKQKOTIGDI-----LPN*TDKKR-----AAPEKRE--QF-KQPR-SC-R-AS--R-DLK-N-V-----Q--Q--Q--DPRSQF

WFC

B

Oct-1 (743)

Oct-2 (467)

Pit-1 (291)

unc-86 (467)

POU DOMAIN

C

Oct-1 BRKKRTSIEENIVALEKSPLENOKPTSEETITNADLNKEEVIIRVUCNRQKQEKRIKPPSS

Oct-2 BRKKRTSIEENIVALEKSPLENOKPTSEETITNADLNKEEVIIRVUCNRQKQEKRIKPPSS

Pit-1 BRKKRTSIEENIVALEKSPLENOKPTSEETITNADLNKEEVIIRVUCNRQKQEKRIKPPSS

unc-86 BRKKRTSIEENIVALEKSPLENOKPTSEETITNADLNKEEVIIRVUCNRQKQEKRIKPPSS

POU DOMAIN

A B

I

II

III

IV

V

VI

Introduction

1.3.2 Oct-1 and Oct-2A/B have a conserved POU domain which directs DNA binding.

Sequence analysis of human Oct-1 and 2 cDNAs has revealed that the proteins encoded have an approximately 150 amino acid region of very high homology. This region was termed the POU domain, since an homologous region also occurs in the rat pituitary specific transcription factor Pit-1 (regulates growth hormone and prolactin gene expression) and the C. elegans gene unc-86 (involved in neural development; it was identified by genetic analysis, and its biochemical function is unknown) (POU = Pit, Oct, unc) (Herr et al (1988)). A POU domain sequence alignment is shown in figure 4A, and the location of the POU domain within the proteins in figure 4B.

Oct-1 and 2 show 87% homology within the POU domain. Between all four proteins there is 37% identity within this region, although this is higher for particular pairs (Oct-2:unc-86 = 42%, Oct-1:Pit-1 = 52%). There is little homology outside this region, except between Oct-1 and Oct-2 which share features such as Q and S/T rich regions, and there is some homology between the N termini of Oct-1 and Oct-2B. These similarities will be discussed in a later section.

The POU domain can be divided into 2 subdomains: the POU homeo box (a region at the C-terminal end of the homeo box, known as the WFC sub region, is particularly well conserved) and the POU specific box (which can be divided into A and B sub-regions on the basis of a higher degree of sequence

Introduction

similarity). Between the approximately 60 amino acid subdomains is a region of sequence dissimilarity (approximately 20 amino acids, known as the non-conserved linker). The POU homeo box shows about 30% homology to the antennapedia homeo box (Garcia-Blanco et al (1989)). Homeo boxes were first identified as a conserved region between Drosophila homeotic proteins which are involved in segmental development. Related homeobox proteins have been found in mammals, amphibians and elsewhere. An alignment of the Oct-1 and Oct-2 homeo boxes to the antennapedia homeo box (which is considered to be the consensus Drosophila homeo box) is shown in figure 4C. The Drosophila homeo box has been shown to be a DNA binding motif, and contains a helix turn helix motif (helices 2 and 3 in figure 4C) similar to that seen in bacterial repressor proteins. Bacterial repressor proteins (lambda, 434 and trp) have been analysed by X-ray crystallography and the residues involved in DNA binding determined from protein:DNA co-crystals. A cluster of basic residues (K or R) following helix 3 (II on figure 4C) in homeoboxes is well conserved in the 434 repressor, where it contacts DNA. Similarly, a cluster of basic residues before helix 1 (I on figure 4C) in homeo boxes is conserved in the lambda repressor, where it contacts DNA. The residues marked A and B (see figure 4C) in helix 1 are well conserved at equivalent positions in the bacterial repressors where they have been shown to contact DNA. The basic regions I and II (RRRKEKR: RRRKKR) are conserved at the underlined positions

Introduction

in more than 95% of homeo boxes, including the Oct-1 and 2 homeo boxes. Consequently, by homology to other homeo boxes, and by homology to the bacterial repressors, the homeo box of Oct-1 and 2 is predicted to be the DNA binding domain.

The POU domain, containing the homeo box, has been shown to be responsible for DNA binding. Oct-1 N terminal deletions up to the POU domain have no effect on DNA binding (Sturm and Herr (1988)). Oct-1 C terminal deletions have no effect on binding up to the POU domain (Sturm et al (1988A)). However, deletions into the POU domain abolish binding. Mutations in POU specific box A and B subregions abolish binding, as do mutations in homeo box helix 3. However, addition of six A residues to the non-conserved linker has no effect on DNA binding. The same mutations affect binding to two different motifs (SV40 octamer 1, and HSV alpha gene TAATGARAT motif) in a similar way. As already noted, Drosophila homeo box proteins are known to be DNA binding proteins and contain the homeo box in the absence of the POU specific box. Consequently the function of the POU specific box is not clear. It may not contact the DNA directly, but could, for example, stabilise interactions between the homeo box and DNA. The POU homeo box alone has a low level of DNA binding activity (Verrijzer et al (1990)), which supports the suggestion that the the POU-specific box stabilises interaction with DNA. The homeo box alone has altered DNA binding specificity, and so the POU specific box may also have a role in determining the specificity of binding. A

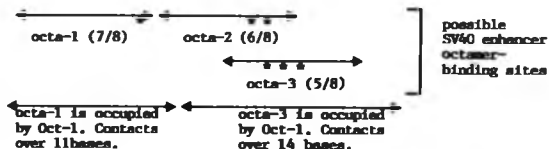
Figure 5.

Oct-1 binds to degenerate octamer motifs by association with flanking sequences.

- (A) Part of the sequence of the SV40 enhancer *aph* motifs is shown, and 3 possible octamer binding sites are indicated. * represents a base which is different to the consensus octamer motif (ATTGCGAT). The occupancy of octa-1 and octa-3 sites by Oct-1 is illustrated. The mutations indicated were made in the octa-3 site, and Oct-1 binding assayed. (See Sturm et al (1987), Baumrucker et al (1988)).
- (B) The SV40 octa-3 site is 14 bases long. If the HSV alpha gene Oct-1 site is aligned to the SV40 octa-1 site over the corresponding 14 bases, they match at only 4 of the 14 positions.

[A]

TATGCAAAGCATGCAATCTCAATTAG



W.T. octa-3
(~differences
to consensus)

OCTAMER
A T G C A T C T C A A T T A G
* * *

Oct-1
BINDING:
+

W.T. octa-3 with
mutated flanks
(underlined).

T A G C A T C T C A A T T A G
* * *

-

Consensus octamer
octa-3.

A T G C A T G C A A A T T A G

+

Consensus octamer
octa-3 with
mutated flanks.

T A G C A T G C A A A T T A G

+

[B]

HSV alpha gene motif G G T A A T G A G A T G C C 4/8 octamer
SV40 octa-1 site A T G C T T T G C A T A C T 7/8 octamer

* = difference to consensus octamer
| = conserved base

Introduction

similar deletion series, and mutations in the POU domain demonstrate that the Oct-2 POU domain is required for DNA binding (Ko et al (1988), Muller-Immergluck et al (1990)).

As already noted, Oct-1/Oct-2 (via the POU domain) bind remarkably degenerate octamer motifs including the HSV alpha gene TAATGARAT motif, the Ig heptamer motif, two sites in the SV40 enhancer, sites in the 7SK RNA gene promoter, and so on. The SV40 enhancer octamer sites have been analysed to determine the sequence requirements for binding to degenerate octamer motifs (Sturm et al (1987), Baumruker et al (1988)). As shown in figure 5A, the SV40 enhancer contains three potential octamer binding sites (Octa 1 to 3). DMS and DEPC modification protection assays and oligonucleotide mutagenesis have been used to show that it is in fact the Octa 1 (7/8 match) and Octa 3 (5/8 match) sites which are occupied by Oct-1. The 6/8 match Octa 2 site (which overlaps the Octa 3 site) is not occupied. Oct-1 contacts over a region of 11 bases to associate with the Octa 1 site, and over a region of 14 bases to associate with the Octa 3 site. This suggests that flanking sequences may be important to allow binding to degenerate octamer motifs. If the Octa 3 site is mutated to contain the consensus octamer (see figure 5A) then the mutation AT to TA at the 5' end of the binding site has no effect on Oct-1 binding, whereas this mutation abolishes binding to the wild-type Octa 3 site. The stabilisation of binding to degenerate motifs by flanking sequences seems to hold true for other sites, where the

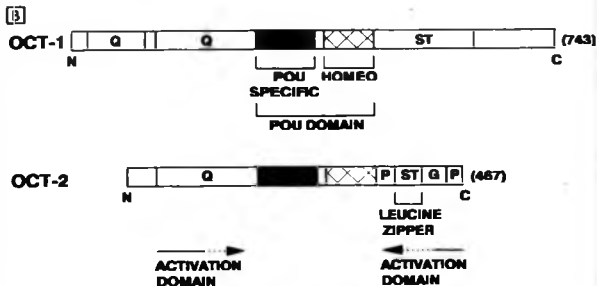
Figure 6.

Homology between Oct-1 and Oct-2 outside the POU domain.

- (A) A sequence alignment of the N terminal ends of human Oct-1 and mouse Oct-2B. | = a perfect match, and : a conservative match (I:L or S:T). (See Hatzopoulos et al (1990)).
- (B) Apart from the POU domain and the N terminal ends of Oct-1 and Oct-2B, homology between Oct-1 and Oct-2 consists of regions rich in particular amino acids. Amino acid rich regions are illustrated here. Also shown is the location of the putative Oct-2 leucine zipper and the Oct-2 activation domains (represented by arrows). The solid part of the arrow represents the region which contributes most to transcriptional activation. (See Sturm et al (1988), Clerc et al (1988), Scheidereit et al (1988), Muller et al (1988), Tanaka and Herr (1990) and Gerster et al (1990)).

616 678
OCT-1 SLPNGTSLGSLSPALNSKNSTVATIQALASGGLPLTSLDQNLVF ANAGCAPNTVATAPLPLN
Oct-2B GSNVGLSSGLSPALNSNPLATTQALASGGLPLTSLDQSNVFLGAGCAAPGSPSLVTSPLPLN
451 516

679 743
OCT-1 PQNLSLLTNP VSLVSAAMASAGSAPVASIHAISTSAESTQNSLFTVASASGAASTTTTA SKAE
Oct-2B HTGLPLLSAPPQVGLVSAAMAAVASTSSKSPGLSSSSSSSSSSSTCSDAVATITPGPGPGPEACKAE
517 583



Introduction

dependence on flanking sequences is linked to the degree of homology of the core octamer motif to the consensus (Baumrucker et al (1988)). The extent of the degeneracy of sites to which Oct-1 is able to bind is illustrated in figure 5B. Oct-1 binds the HSV alpha gene TAATGARAT motif (4/8 octamer) and the SV40 enhancer Octa 1 site. The longest region over which Oct-1 contacts is 14 bases (SV40 Octa 3 site). Out of these 14 bases the SV40 octa 1 and TAATGARAT motifs have only 4 bases in common.

1.3.3 Other features determined by examination of the sequences of cloned Oct-1 and Oct-2A/B.

As mentioned earlier, alternate splicing causes the last 12 amino acids of Oct-2A to be replaced by 132 new residues in Oct-2B. The N terminal 127 residues of Oct-2B show 58% homology to the N terminal 121 residues of Oct-1 (see figure 6A). Within this region a stretch of 81 residues is 79% homologous. This extensive homology suggests a similar function.

Apart from the POU domain, and the homology described above, any homology between Oct-1 and Oct-2 is a consequence of regions rich in particular amino acids. The N terminal domains are Q rich (residues 22 to 268 in Oct-1 are 26% Q) and parts of the C terminal domains are S/T rich (residues 441 to 560 in Oct-1 are 50% S/T). This arrangement is reminiscent of the ubiquitous mammalian transcription factor, Sp1 (reviewed in Mitchell and Tijan (1989)). Oct-2 contains P

Introduction

rich regions in the C terminal domain. Proline rich regions have been identified as activation domains in the transcription factor CTF/NF-1 (reviewed in Mitchell and Tijan (1989)). As will be described later, these regions have been shown to be important for transcriptional activation.

Oct-2 contains a potential leucine zipper. Four leucine residues are each separated by exactly seven other residues. Such structures are believed to form an amphipathic alpha helix, which can align and pair via the hydrophobic leucine residues to another leucine zipper, leading to dimerisation or interaction with other proteins (reviewed in Busch and Sassone-Corsi (1990)). This motif is not seen in Oct-1, and as will be described later is not required for transcriptional activation by Oct-2 in HeLa cells. However, it could still be required in B cells. In immunoglobulin promoters/enhancers the octamer motif is part of an array of factor binding sites, and perhaps the leucine zipper allows Oct-2 to interact with other factors.

Introduction

1.3.4 Transcriptional activation by cloned Oct-1 and Oct-2A.

Three recent papers have analysed the regions of Oct-1 and Oct-2A required for transcriptional activation by the cloned factors in transfected HeLa cells (Tanaka and Herr (1990), Muller-Immergluck *et al* (1990), Gerster *et al* (1990)).

Tanaka and Herr (1990) used the SV40 enhancer sph motifs upstream of the the β -globin TATA box, upstream of the β -globin CCAAT-TATA boxes, and 2.2 kb downstream of the β -globin transcription start site as targets for activation by co-transfected Oct-1/Oct-2 in HeLa cells. Six copies of the Sph motif were used in each case, with various mutations (oct⁺sph⁺, oct⁺sph⁻, oct⁻sph⁺, oct⁻sph⁻) as described in section 1.2.2. Oct-2 activates from oct⁺sph⁻ in either upstream position, whereas Oct-1 does not. As a control, there is no activation from oct⁻sph⁻. These results are consistent with those obtained by transfection into HeLa cells (expressing Oct-1) and B cells (expressing Oct-1 and Oct-2) described in section 1.2.2. However, neither factor activates from oct⁺sph⁻ when it is located downstream of the promoter. Oct⁺sph⁻ is active in the downstream position in transfected B cells (section 1.2.2), although with 12 fold less activity than when located upstream of the TATA box. Consequently in B cells there may be another octamer-binding factor, or modification of Oct-2, or interaction of Oct-2 with another factor to allow the octamer motif to operate from the downstream position in B cells.

The oct⁺sph⁺ promoter displays reduced activity if

Introduction

co-transfected with Oct-1 or Oct-2 in HeLa cells. This effect is not seen with the oct⁻sph⁺ promoter. So, Oct-1/Oct-2A probably displace sph factors, leading to reduced transcription. Mutations in the POU (DNA binding) domain abolish the repression effect, and factors with deletions outside the POU domain have the same repressive effect. Consequently the effect is probably due to DNA binding, rather than an inhibitory domain. This repression was used to normalise the amount of Oct binding activity present in the nucleus in subsequent assays where deleted factors were used to locate activation domains.

Oct-1 and Oct-2A were divided into N terminal (N), POU (P) and C terminal (C) domains, and chimeras made. (Oct-1 C, P + Oct-2 N) and (Oct-1 C + Oct-2 P, N) did not activate transcription, whereas (Oct-2 C, P + Oct-1 N) and (Oct-2 C + Oct-1 P, N) did activate transcription from the octamer upstream of the β -globin promoter in HeLa cells. Consequently the ability of Oct-2 to activate this promoter resides in the C terminal domain, and the N terminal and POU domains are interchangeable between the two factors.

Oct-2 N and C terminal deletions have been used to map the activation domains. Deletion of the first 100 or so residues of the N terminal domain has no effect on the ability to activate transcription from the octamer motif. However, deletions which extend into the Q rich regions (see figure 6B) decrease activation. C terminal deletions also decrease activation. Consequently Oct-2 has two interdependent

Introduction

activation domains. Mutations in the leucine zipper residues (see section 1.3.3) have no effect on activation by Oct-2 in this assay system. Similar results have been obtained by two other groups (Muller-Immergluck et al (1990), Gerster et al (1990)). However, this sequence motif could be important in other cell types.

Phosphorylation of Oct-1, Oct-2 and chimeric constructs was analysed by comparison of in vitro translation products with immunoprecipitated products from transfected HeLa cell which had been labelled with ^{35}S -methionine or ^{32}P , and with these products after treatment with phosphatase. Oct-1 and Oct-2 are both heavily phosphorylated and Oct-2 and active chimeras (but not Oct-1 and inactive chimeras) display an electrophoretic mobility shift on phosphorylation. Consequently the phosphorylated forms displaying the mobility shift are probably involved in the activation of transcription. Oct-2 activation domain deletions decrease the overall level of phosphorylation, but the electrophoretically retarded phospho-proteins still occur indicating that the activation domains induce phosphorylation elsewhere in the protein.

Essentially equivalent results, indicating that Oct-2 has interdependent N and C terminal activation domains were obtained by Gerster et al (1990). Here the domains were mapped more precisely by internal deletions. The N terminal activation domain includes the Q rich region, and an adjacent P rich domain contributes to some extent. The first

Introduction

approximately 100 amino acids at the N terminus are not required for activation. Most of the C terminal domain contributes to activation, with the P rich region being most important. The Oct-2 activation domains are illustrated in figure 6B.

Conflicting results concerning the Oct-2 activation domains have been reported by Muller-Immergluck et al (1990). Their Oct-2 deletion analysis in a similar assay system suggested that N and C terminal deletions alone have little effect on activation potential. Only N and C terminal deletions together give reduced activity. This suggests that activation domains are not interdependent, but functionally redundant. An interesting finding described in this paper is that the activation potential of deleted Oct-2 varies according to the position of the octamer relative to the TATA box. With the octamer motif 5 bases from the the TATA box deletion of the C terminus has no effect on the activation potential of Oct-2, whereas with the octamer motif 20 bases from the the TATA box there is a 2 fold decrease in transcription (Gerster et al (1990) report a similar decrease in activation with the same promoter and a similar Oct-2 deletion, but since the deletion causes a 60-100 fold increase in binding activity, the deleted Oct-2 is essentially inactive with respect to the full length Oct-2. However, a direct comparison of the ability to activate from the 2 promoters is still valid). Also, with the octamer 5 bases from the TATA box the POU domain alone is able to activate transcription to some

Introduction

extent. Consequently, the spatial arrangement of the promoter may be important for activation by Oct-2.

As mentioned above, Oct-2 deletions are expressed at drastically different levels compared to the full length protein (up to 100 times higher). This could account for the contrasting results reported by Muller-Immergluck et al (1990). Tanaka and Herr (1990) normalised activation potential against the amount of binding activity in the nucleus using the Sph repression assay, and Gerster et al (1990) against the amount of binding activity seen in band shift assays, after demonstrating that the deletions were able to enter the nucleus (since co-transfected deletions, unable to activate transcription, were able to inhibit activation by full length Oct-2). Perhaps the normalisation of activation potential to the amount of nuclear binding activity was not as accurate in the assays performed by Muller-Immergluck et al (1990).

A by-product of the normalisation used in these assays was to show that Oct-2 deletions reached the nucleus. Consequently the nuclear localisation signals do not reside in the N or C terminal domains of Oct-2.

1.3.5 The Oct-1 POU domain interacts with VP16.

As already noted VP16 forms a complex with the Oct-1:DNA complex. However, purified VP16 and Oct-1 (assayed by band shift assays) do not form this complex with the HSV alpha gene TAATGARAT motif (Kristie et al (1989)). A third cellular

Introduction

component (found in insect and HeLa cells) is required. This 3 component complex protects regions flanking the octamer, and mutations in these sequences abolish complex formation (although the flanking sequence requirements must be flexible, since VP16 can transactivate from the TAATGARAT motif, the H2B octamer, the U3 snRNA octamer and the SV40 enhancer octamer - see section 1.2.1). A further complex is formed by interaction with another factor found in HeLa cells (but not insect cells). The footprint pattern of this complex is identical to the previous complex, suggesting that formation involves only protein-protein interactions (and not protein-DNA interactions). Formation of these complexes requires only the POU domain of Oct-1. Oct-2 forms similar complexes, but only at very high levels of VP16.

Stern et al (1989) have shown that the POU homeo box is required for interaction with VP16. The Oct-1 and Oct-2 homeoboxes differ at 7 out of 60 positions. If the residues at three of these positions in Oct-1 are mutated to those seen in Oct-2 there is no effect on DNA binding, but formation of the complex with VP16 is disrupted.

The interaction of Oct-1 with other proteins has obvious implications in explaining how Oct-1 regulates diversely expressed genes, for example the possibility that a cellular analogue of VP16 is involved in the cell-cycle regulation of histone H2B genes has already been discussed. The inability of the Oct-2 POU domain to interact efficiently with VP16 is

Introduction

probably part of the explanation for how two factors binding the same element can differentially activate transcription.

1.3.6 The POU domain is sufficient to stimulate DNA replication.

As already described the octamer motif is conserved in the adenovirus origin of replication, and purified Oct-1 stimulates adenovirus replication in vitro from the octamer motif (section 1.1.7). Over-expression of cloned Oct-1 and Oct-2A in vivo stimulates adenovirus replication in an octamer-dependent manner (Verrijzer et al (1990)). Deletion of Oct-1 indicates that only the POU domain is required to stimulate replication. Verrijzer et al (1990) report that the POU homeo box alone has a low level of DNA binding activity. The Pou homeo box alone inhibits DNA replication (presumably by blocking the binding of endogenous Oct-1), suggesting that the POU specific box is the region required for the stimulation of DNA replication.

Introduction

1.4 Other POU domain and octamer-binding proteins.

1.4.1 Homeo box-containing genes form multi-gene families in Drosophila and vertebrates.

Homeotic mutations transform certain parts or an entire body segment into the corresponding structures of another body segment. In Drosophila genetic analysis lead to the identification of homeotic genes specifying segmental identity (for review see Gehring (1987)). Structural analysis identified a conserved DNA segment (approximately 180bp) between homeotic genes, known as the homeo box. Using the homeo box as probe a battery of homeotic genes have been identified in Drosophila, and in vertebrates. The POU domain of Oct-1/Oct-2 contains a homeo related box with approximately 30% identity to the Drosophila antennapedia homeobox (considered to be the consensus homeo box, see section 1.3.2).

As described in section 1.3.2 the homeo box is related to the helix-turn-helix DNA binding motif of bacterial repressor proteins. Consequently homeo box proteins were postulated to be DNA binding proteins, and probably transcription factors. Using similar techniques to those described for Oct factors, homeo box proteins have been shown to bind DNA via the homeo box, and to activate transcription (reviewed in Levine and Roey (1988), Biggin and Tijan (1989)). In a manner reminiscent of Oct-1 and Oct-2 the *en*, *eve*, *ftz*, *prd* and *Ubx* homeo box proteins of Drosophila recognise some DNA sequences with similar affinity, whereas others (eg. *bicoid*) seem to

Introduction

have distinct binding affinities. In transfection assays introduction of Drosophila homeo box cDNAs along with target promoters/enhancers has shown that Drosophila homeo box proteins are transcription factors. Consequently many of the questions which have been discussed concerning how Oct-1 and Oct-2 are able to differentially activate transcription from the same promoter element also arise for homeo box proteins in general. As a result studies on octamer binding transcription factors have relevance to differential gene activation, and hence specification of segmental identity by homeo box proteins. For example, the discovery that Oct-1 interacts with VP16 via its homeo box, and that 3 residue changes (to those seen in Oct-2) prevent this interaction may indicate a mechanism for differential gene activation applicable to homeo box proteins in general (see section 1.3.5).

Using Drosophila homeo boxes as probe, families of homeo box proteins have also been identified in vertebrates (for reviews see Dressler and Gruss (1988), Dressler (1989), Wright et al (1989)). In view of this it might be supposed that families of POU domain/octamer-binding proteins occur in vertebrates. As will be described below, apart from the already isolated Pit-1 (from rat) and unc-86 (from C. elegans) identification of Oct-1/Oct-2 has lead to the discovery of an array Oct/POU proteins (mostly in mammals) by band shift assays and screening with POU domain probes.

Introduction

1.4.2 The unc-86 gene of C. elegans.

In the nematode C. elegans mutations in the unc-86 gene cause specific defects in the nervous system. The unc-86 gene was cloned by genetic mapping and chromosome walking (Finney et al (1988)). Open reading frames and consensus splice sites were used to determine the sequence of the transcript from the gene and the protein encoded. As described in section 1.3.2 the protein encoded contains a POU domain (POU = Pit, Oct, unc). By analogy to Oct-1/Oct-2 and Pit-1, the unc-86 product is presumably a transcription factor, however the target genes for this putative transcription are unknown. The loss of unc-86 function causes certain blast cells to not express their normal fate (but to display the characteristics of their mother) and a particular neuron displays its normal characteristics early in development, but subsequently fails to mature and remains in the quiescent state.

1.4.3 The pituitary-specific transcription factor, Pit-1

Growth hormone (GH) is specifically expressed in somatotrophs of the anterior pituitary. Prolactin (Prl) is expressed in lactotrophs, most of which are derived from GH producing precursors (reviewed in Karin et al (1990)). A protein GHF-1 (said to be somatotroph specific) was identified and bound to two adjacent sites in the GH gene promoter which are important for in vitro and in vivo expression of the gene. The purified GHF-1 protein stimulated the GH promoter in conjunction with extract from cells which

Introduction

do not express GH or GHF-1 (HeLa cell extract). GHF-1 was reported to not bind the Prl promoter. However, another group identified a rat factor, Pit-1 (later shown to be identical to GHF-1) which bound the same sites in the GH promoter, and also bound (possibly with lower affinity) to the prolactin promoter. Pit-1 was reported to be expressed in GH and Prl producing cells (see Karin et al (1990), Nelson et al (1988) and references therein).

GHF-1 (Bodner et al (1988)) and Pit-1 (Ingraham et al (1988)) have been cloned, and are identical. GHF-1 was cloned by screening with probes derived from the partial amino acid sequence of the purified protein, and Pit-1 by screening an expression library with the binding site. The 291 amino acid protein (henceforth referred to as Pit-1) contains a POU domain at the C terminal end of the protein. In a similar manner to that described for Oct-1/Oct-2, the Pit-1 POU domain has been shown to be required for DNA binding. The Pit-1 activation domain resides in the N terminal half of the protein. The first 72 amino acids of the protein (which are S/T rich) have been shown by deletion analysis to be required for transcriptional activation. Pit-1 binds multiple sites in the Prl and GH gene promoters, which are required for transcriptional activity in vivo and in vitro (Nelson et al (1988)). The consensus binding site is $A/TATNCAT$, which is clearly related to the octamer motif (ATTGCAT). Cross-competition with GH and Prl promoter binding sites in in vitro transcription assays indicates that a common factor

Introduction

activates both genes (Nelson et al (1988)) and cloned Pit-1 activates expression from both GH and Prl promoters on co-transfection into HeLa cells (Ingraham et al (1988)). Stable transfectants expressing cloned Pit-1 at a level 10 fold lower than physiological activate both GH and Prl promoters. Activation from the GH promoter is much less than from the Prl promoter, even though bacterially expressed Pit-1 has at least as high affinity for the GH promoter Pit-1 binding sites (Mangalam et al (1989)). So, the evidence that Pit-1 (GHF-1) activates two genes with different patterns of expression is convincing. This begs the recurring question 'How can the same factor specify differential gene expression?' Modification or interaction with other factors are possibilities. Other cis-acting sequences occur in GH and Prl promoters, and binding to these sites has been observed. As well as being expressed in somatotrophs where it stimulates GH (and not Prl) gene expression and lactotrophs where it stimulates Prl (and not GH) gene expression, Pit-1 protein is also present in the thyrotrophs of the anterior pituitary where it stimulates neither GH or Prl gene expression (Simmons et al (1990)). Pit-1 mRNA is also detected in the other two cell types of the anterior pituitary (gonadotrophs and corticotrophs), however Pit-1 protein is not present. This suggests that cell-specific translation of the Pit-1 mRNA occurs.

A rat prolactin promoter Pit-1 binding site was inserted in single copy in front of the Prl gene TATA box, and a series

Introduction

of mutations analysed for Pit-1 binding affinity and transcriptional activation in Pit-1 expressing cells (Elsholtz et al (1990)). The in vitro binding affinity does not necessarily reflect the ability of a Pit-1 site to activate transcription, since some low affinity binding sites show significant ability to activate transcription in vivo. Oct-2 binds the wild-type Pit-1 binding site, but the promoter is not active in B cells. However, conversion of the core of the Pit-1 site to a perfect octamer (which is a higher affinity Oct-2 binding site) creates a promoter which is active in B cells. Pit-1 shows significant, but lower, affinity for the consensus octamer promoter, but this promoter is not active in Pit-1 expressing cells. The explanation here could be that, in vivo, auxiliary factors stabilise the interaction of Pit-1 with some low affinity sites, and also that auxiliary factors are able to destabilise the interaction of Pit-1 with the consensus octamer promoter (and destabilise the interaction of Oct-2 with the Pit-1 site).

The Pit-1 gene has been cloned, and the factors required for its expression analysed (McCormick et al (1990)). The promoter contains two CREB (a cAMP responsive transcription factor) binding sites, and expression is cAMP responsive. The promoter also contains a Pit-1 binding site. Mutation of this site reduces expression of the gene, and expression can be stimulated by cloned Pit-1. Consequently positive autoregulation takes place. Pit-1 binding site mutants are

Introduction

still expressed in a cell-type specific manner, and so other factors must be involved in the the cell-type specific expression of the gene. Similar results have been obtained by a second group (Chen et al (1990)), except that a second Pit-1 binding site has been identified downstream of the transcription start site. Mutation of this site causes a Pit-1-dependent increase in expression, so Pit-1 binding to this site is inhibitory (perhaps by halting transcription, rather than by inhibition of the initiation of transcription). It is suggested that Pit-1 stimulates its own expression, but that the second inhibitory site limits stimulation to a lower level than other target genes.

Introduction

1.4.4 A family of POU domain genes in the mammalian brain.

Degenerate oligonucleotides for all possible codons of two conserved nine amino acid stretches at either end of the POU domain were used to clone the POU domain of four new POU domain-containing genes by the polymerase chain reaction (He et al (1989)). Brn-1 and Brn-2 were isolated from human brain, Brn-3 from rat brain, and Tst-1 from rat testes. Brn-1, Brn-2 and Tst-1 are about 94% similar to each other throughout the POU domain, and 85% similar to Oct-1/Oct-2 in the POU specific box, and 64% similar to Oct-1/Oct-2 in the homeo box. Brn-3 is much more closely related to unc-86 than any of the other POU domain genes being 79% similar within the POU domain.

In situ hybridisation was used to analyse the expression of these POU domain genes (and Pit-1, Oct-1, Oct-2) in the adult and developing mammalian brain. Expression of these four new POU domain genes is both temporally and spatially restricted in the mammalian brain, with each gene displaying a unique pattern of expression. As already noted (see section 1.3.1) Oct-2 is widely expressed in the developing brain, and displays a restricted pattern of expression in the adult brain. Pit-1 is expressed early in brain development, with transcripts disappearing, and then reappearing exclusively in the anterior pituitary. Oct-1 expression is widespread in the developing brain, but surprisingly in the adult brain is highly restricted to particular regions.

Introduction

1.4.5 A family of octamer-binding proteins in mouse embryogenesis.

Band shift assays with an octamer-containing probe (from the IgH enhancer) have been used to analyse mouse tissues and stages of embryogenesis for the presence of octamer-binding proteins (Scholer et al (1989)). Ten distinct octamer binding activities were identified (including Oct-1 and Oct-2) and termed Oct-1 to Oct-10. These factors have distinct distributions. Oct-3 is expressed in the embryo and adult brain, and when present is accompanied by Oct-2 and Oct-7 (these factors occur without Oct-3). As already noted, Oct-2 transcripts can be detected in tissues other than lymphoid (see section 1.3.1). In line with this, Oct-2 like complexes are seen with extract from embryo, brain, spleen and kidney. However, these complexes can be distinguished from B cell Oct-2 on the basis of stability at increased temperature and affinity for different probes. These factors could represent modifications of B cell Oct-2, or distinct proteins.

So far, most interest has centred on Oct-4 and Oct-5. Oct-4 is in both male and female primordial germ cells. Oocytes (and not sperm) and embryonic stem cells contain both Oct-4 and Oct-5. Neither factor is present in the post-implantation embryo. Oct-4 and Oct-5 are, hence, maternally expressed octamer binding proteins. F9 and D3 embryonic stem cell lines contain Oct-4 and 5 (and not Oct-2). The octamer motif placed upstream of a reporter gene forms a strong enhancer in these stem cells (Scholer et al (1989A)). On differentiation of

Introduction

stem cells, both Oct-4/Oct-5 levels and enhancer activity decrease, suggesting that Oct-4 and/or Oct-5 are transcriptional activators. A similar reporter gene is expressed in an octamer-dependent manner in the inner cell mass of the blastocysts (source of stem cells) in transgenic mice. As described in section 1.1.1 contrasting results have been obtained by another group (Lenardo et al (1989). Here, a more complete IgH enhancer was placed upstream of a different reporter gene. Octamer mutations caused the enhancer to become active in F9 cells, suggesting the existence of a repressive octamer-binding factor. A novel octamer-binding factor (NFA-3) was identified, and levels decreased on differentiation of F9 cells. This was accompanied by activation of the enhancer. The factor NFA-3 is probably equivalent to Oct-4, since the Oct-5 binding activity is weaker and may not have been detected. A possible explanation for the contradictory results is that the context of the octamer determines the effect of factor binding. As described in section 1.3.4, transfection of Oct-1 and Oct-2 can activate transcription from the SV40 enhancer in HeLa cells if the sph motif inactivated, but decreases the activity of the SV40 enhancer if the sph motif is intact (presumably by blocking sph factor binding). In a particular promoter arrangement, Oct factor binding in F9 cells may inhibit the formation of a transcription initiation complex.

Three groups have reported the cloning of a novel octamer-binding/POU domain factor from embryonic stem cells (Scholer

Introduction

et al (1990), Rosner et al (1990), Okamoto et al (1990)). Rosner et al (1990) report a 352 amino acid open reading frame. Scholer et al (1990) report a 324 amino acid ORF. The cDNAs are identical in sequence except that the Scholer et al (1990) sequence is shorter at the 5' end and contains an extra A base close to the 5' end. The extra A creates an in frame first ATG, causing the predicted protein to be shorter by 28 amino acids than that reported by Rosner et al (1990). Relative to the sequence reported by Rosner et al (1990), the sequence of Okamoto et al (1990) lacks a G residue in the C terminal third of the protein, resulting in a different reading frame at the 3' end. Since the sequence of Rosner et al (1990) matches one sequence at the 5' end, and the other sequence at the 3' end it is probably the correct sequence for the factor Oct-4 (called Oct-3 by Rosner et al (1990) and Okamoto et al (1990)). Oct-4 is about 75% similar to Oct-1/Oct-2 within the POU specific box, and 58% similar within the homeo box.

Scholer et al (1990) report that the in vitro translation product of their clone gives two bands in a band shift assay, of equivalent mobility to Oct-4 and Oct-5 from F9 cells. It is possible that both proteins are translated from the same message by use of alternate initiation codons. Rosner et al (1990) report a similar effect, but do not detect Oct-5 in extract from stem cells. Co-transfection of Oct-4 with an octamer-containing reporter gene in HeLa cells activates transcription from the gene, showing that Oct-4 is a

Introduction

transcription factor. Rosner et al (1990) were able to detect Oct-4 message in stem cell lines (decreases on differentiation), oocytes, pre-implantation embryos, and soon after no expression was detected. This matches the band shift data (Scholer et al (1989)), except that no Oct-4 protein is detected in the post-implantation embryo (expression of the message at this stage is restricted, and so the protein is probably of too low abundance to be detected in extract from whole embryos). However, in contrast to the band shift data, Oct-4 message is detected in testis. RNase protection assays with ovary RNA give the fully protected species, and a smaller protected band, indicating that two related messages are present. Only the fully protected band is seen with stem cell RNA.

1.4.6 Other examples.

The sea urchin sperm specific histone H2B gene promoter contains two adjacent inverted CCAAT boxes flanked at either side by an octamer motif (Barberis et al (1987), Barberis et al (1989)). In spermatocytes (expressing the gene) both CCAAT boxes are occupied, whereas in the embryo (non-expressing) only the CCAAT box most proximal to the TATA box is occupied, and by a distinct CCAAT binding factor (CCAAT displacement protein) which seems to block any interaction at the adjacent site. Sea urchin has an apparently ubiquitous octamer-binding protein (presumably Oct-1) of similar mobility to mammalian Oct-1 in band shift assays, and of

Introduction

similar molecular weight to mammalian Oct-1. This protein occupies the octamer site in embryos. In testis a second, more abundant and testis specific octamer binding protein is present. This protein (Oct-T) is slightly smaller than Oct-1. It is suggested that both the CCAAT displacement protein (as a repressor in non-expressing cells) and the testis specific octamer binding protein (as an activator in expressing cells) contribute to the testis specific expression of this H2B gene.

In Drosophila the dopa decarboxylase gene (ddc) is expressed in selected dopaminergic neurons. A POU domain protein binding to a ddc gene promoter sequence element, which is necessary for the expression of the gene, has been cloned (Johnson and Hirsh (1990)). As yet, there is no direct evidence that this protein regulates the expression of the ddc gene. As described earlier in this section, several mammalian POU domain proteins with unique and restricted patterns of expression in the nervous system have been identified. However, this is the first POU domain protein which has been associated with a specific function in the nervous system.

A mammalian octamer-binding protein, apparently specific to malignant melanoma cells has been reported (Cox et al (1988)).

Introduction

1.5 Project aims.

In view of the intrinsic interest of the 'octamer paradox' the project began by trying to clone X. laevis octamer binding proteins using a probe from the histone H2B gene promoter to screen expression libraries. When the sequence of human Oct-1 and 2 was reported, the X. laevis homologue of Oct-1 was cloned for the following reasons. Firstly, to make a comparison of Oct-1 between species. Such comparisons can lead to the identification of functionally important regions by the extent of homology. Secondly, it was thought that the oocyte micro-injection system might provide a transcriptional assay system for the activity of the octamer motif and octamer binding proteins. Thirdly, the POU domain had been shown to contain a homeo box. As noted in the previous section, homeo box containing genes in Drosophila and vertebrates occur in large families and are developmentally important. As Xenopus is an organism amenable to the study of development it was decided to see if a family of octamer-binding proteins existed in Xenopus. Some means of identifying and eliminating Oct-1 in binding assays was considered to be appropriate.

Materials

CHAPTER 2: MATERIALS.

2.1 General reagents.

Enzymes were obtained from the following suppliers: restriction endonucleases were from Amersham International (U.K.), Northumbria biologicals limited (NBL) and Bethesda research laboratories, Maryland U.S.A. (BRL); E. coli DNA polymerase I, T4 DNA ligase and T4 polynucleotide kinase were obtained from Amersham International; T3 and T7 RNA polymerases were from New England Biolabs; DNA polymerase Klenow fragment was from NBL; AMV reverse transcriptase was from Life sciences Inc. (U.S.A.).

All radioisotopes were supplied by Amersham International at the following specific activities: α -³²P-dGTP and α -³²P-dCTP, 3000 Ci/mmol; γ -³²P-ATP, 5000 Ci/mmol; α -³²P-UTP, 3000 Ci/mmol; α -³⁵S-dATP, 1000 Ci/mmol; ³⁵S-methionine, >1000 Ci/mmol.

Nitrocellulose filters (Hybond-C) were obtained from Amersham.

Type II agarose (medium EEO) was supplied by Sigma chemical company and low melting point agarose was from FMC Bioproducts.

Materials for bacteriological media were from Difco laboratories (Michigan, U.S.A.) and Oxoid limited (England).

E.coli transfer RNA (type XX) and salmon testis DNA were supplied by Sigma.

Materials

Foetal calf serum was obtained from Gibco, and tissue culture flasks from Falcon.

Acrylamide was supplied by Fisons, and bisacrylamide by Kodak.

X-ray film was from Fuji photo company limited (Japan). Oligonucleotides were synthesised in the department by Gill Scott, using an Applied Biosystems automated synthesiser.

All other chemicals and reagents were from BDH ('Analar' grade) or from Sigma chemical company unless otherwise stated.

2.2 Stock Solutions.

Deionised formamide - Formamide was stirred with Amberlite monobed resin MB-3 or MB-4 (5g per 100 ml of formamide) until the pH was 7, filtered through Whatman No. 1 filter paper and then stored at -20°C.

Filtered formaldehyde - A 40% (w/v) solution of formaldehyde was filtered through Whatman No. 1 paper.

Stock acrylamide solutions were deionised as described for formamide, except for 30 minutes, and then stored in the dark at 4°C.

TE - 10mM Tris.HCl (pH7.5), 1mM EDTA.

10x TBE - 0.9M Tris.borate pH8.3, 20mM EDTA.

20x SSC - 3M NaCl, 0.3M Na Citrate, pH7

50x Denhardt's - 1% (w/v) each of Ficoll, polyvinylpyrrolidone, bovine serum albumin.

Materials

10x MOPS - 0.2M 3-(N-morpholino) propanesulphonic acid,
50mM sodium acetate, 10mM EDTA, pH7
Phosphate buffered saline (PBS) - 0.8% (w/v) NaCl, 0.02%
(w/v) KCl, 0.115% (w/v) Na_2HPO_4 , 0.02% (w/v) KH_2PO_4 ,
1x barth-X - 88mM NaCl, 1.0mM KCl, 2.5mM NaHCO_3 , 15.0mM
Tris.HCl pH7.6, 0.3mM CaNO_3 , 0.41mM CaCl_2 , 0.82mM MgSO_4 .
NAE - 0.3M NaAcetate pH6.5, 1mM EDTA.

2.3 Bacteriological Media.

LB broth - 5g yeast extract, 10g bactotryptone, 10g NaCl
per litre.

H broth - 10g bactotryptone, 8g NaCl per litre.

2x TY - 10g yeast extract, 16g bactotryptone, 8g NaCl per
litre.

NZY broth - 5g NaCl, 2g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 5g yeast extract, 10g
NZ amine per litre.

For plates agar was added at 15g per litre.

For top agar, agar was added at 7g per litre and for top
agarose, agarose was added at 7g per litre.

All media were autoclaved before use.

Antibiotics were used in plates and broth at the following
final concentrations: ampicillin, 100 or 200 ug/ml;
tetracycline, 15 ug/ml.

Materials

2.4 Bacteria, plasmids and phage.

2.4.1 Genotypes of E. coli strains.

JM101 - supE, thi, Δ (lac-proAB), [F', traD36, pro AB, LACiqZDMI5] 17-18 ditto, not tra D36.

MC1061 - F⁻, ara D139, Δ (ara, leu)7696, Δ lac Y74, gal U⁻, gal K⁻, hsr⁻, ham⁺, strA.

Y1090 - [Δ Lac U169, proA⁺, lon, araD139, strA, supF, trpC22::Tn10(pMC9)].

BB4 - supF58, supE44, hadR514 (rk⁻, mk⁻), galK2, galT22, trpR55, metB1, tonA, lambda⁻, D(arg-lac)U169 [F', proAB, lacIqZDMI5, Tn10(tet^r)].

71:18 mutL - Δ (lac-proAB), supE, thi; F' lacI^q Z M15, proA⁺B⁺, Tn10(tet^r), mutL.

TG2 - supE, Δ (lac-proAB)hds, Δ (src-recA) 306::Tn10(tet^r) [F' traD36 proAB⁺ lacI^q lacZ M15 rk⁻ r_m⁻ Rec⁻].

B121(DE3) - F⁻ hds gal, r_B⁻ m_B⁻, lacUV5-T7 gene 1 cloned into int (see Studier and Moffat (1986)).

2.4.2 Plasmid vectors.

pBR322 - General purpose cloning vector (Bolivar (1978)).

pBR325 - pBR322 with chloramphenicol resistance added (Prentki et al (1981)).

pAT153 - High copy variant of pBR322 (Twigg and Sherratt (1980)).

pBluescript SK⁻ - The plasmid generated by automatic excision from lambda ZAP. Has an extensive polylinker, the

Materials

lacZ colour selection system and promoters for both T3 and T7 RNA polymerases (Stratagene, La Jolla, CA, U.S.A).

pET-3 series - Plasmids designed for the high level expression of cloned sequences in E. coli (Rosenberg et al (1987)). 3 vectors with a Bam HI site in each reading frame to make in frame fusions to the first 11 amino acids of gene 10 of phage T7. The fusion gene is transcribed by T7 RNA polymerase, and can therefore be expressed in the E. coli strain BL21(DE3). This strain contains the T7 RNA polymerase gene incorporated into the genome, and under the control of the lacUV5 promoter (inducible with IPTG).

2.4.3 Bacteriophage vectors.

M13mp18 and M13mp19 - bacteriophage M13 vectors designed for sequencing by the dideoxy chain termination method (Messing et al (1981)).

2.4.4 Plasmid and Bacteriophage Recombinants.

pRW23 - X. laevis histone gene cluster Xlhw23 in pBR325 (pers. comm., R. W. Old). Gift of R. W. Old.

pBS - Oct-1* - human Oct-1 cDNA in pBSM13+ (Stratagene) (Sturm et al (1988A)). Gift of W. Herr.

albONE Δ-670 - X. laevis albumin (5' end):H3 histone (3' end) fusion gene in M13mp19 (Old et al (1988)). Gift of A. R. Brooks.

X. laevis mature oocyte cDNA library in lambda ZAP (vector described in Short et al (1988)). Cloned sequences can be

Materials

automatically excised from the phage vector and converted to the plasmid vector pBluescript, for ease of handling. Gift of J. Shuttleworth.

Methods

CHAPTER 3: METHODS.

3.1 Tissue Culture.

3.1.1 Growing *X. laevis* Xtc cells.

X. laevis Xtc cells (Pudney et al (1973)) were grown in Glasgow modified Eagle medium (GMEM) supplemented with 10% (v/v) foetal calf serum, penicillin (10ug/ml) and streptomycin (10ug/ml) at 25°C in 5% carbon dioxide.

The cell line grows as a monolayer and was divided 1 in 3 about every 3 days using the following procedure. The medium was removed and the cells washed in 2.5mls of versene:trypsin mix. The cells were then incubated in 2.5mls of fresh versene:trypsin mix for approximately 2 minutes, whilst agitating the flask to detach the cells from the surface. 30mls of medium was added and the cell suspension distributed between three 75cm² tissue culture flasks.

3.1.2 Growing mouse L cells

Mouse L929 (Ltk-) (Kit et al (1963)) cells were grown exactly as described for *X. laevis* Xtc cells, except at 37°C.

3.1.3 Blocking *X. laevis* Xtc cells at stages of the cell-cycle.

Xtc cells were blocked with hydroxyurea by adding the inhibitor to a final concentration of 5mM in the culture

Methods

medium of subconfluent cells. The cells were incubated with hydroxyurea for 90 minutes, after which time band shift extract was prepared from the cells (see section 3.31.3).

Subconfluent Xtc cells were blocked by serum-starvation by washing twice in GEM and then incubating in fresh unsupplemented GEM for 18 hours. Supplemented medium was then added back and the cells maintained for a further 6 hours, after which time band shift extract was prepared from the cells (see section 3.31.3).

3.2 Obtaining Oocytes, Eggs And Embryos.

3.2.1 Oocytes

Oocytes were obtained by anaesthetising a female Xenopus laevis with MS222 and carrying out a partial ovariectomy. The oocytes were manually stripped from the ovary, washed in 1x Barth-X and maintained in 1x Barth-X at 20°C.

Oocytes were staged according to Dumont (1972).

3.2.2 Eggs and embryos.

Female X. laevis were induced to ovulate by subcutaneous injection with 250 units of serum gonadotrophin (Intervet U.K. Ltd) 48 hours before laying followed by 250 units of chorionic gonadotrophin B (Intervet U.K. Ltd) 16 hours before laying. For 'natural' matings male X. laevis were injected with 250 units of chorionic gonadotrophin 8 hours before laying. For in vitro fertilisation a male X. laevis was killed by injection with 0.5ml of 250mg/ml

Methods

phenobarbitone (suthatal, May and Baker), and the testis removed into 1x barth-X on ice. 50 to 100 eggs layed in 1x barth-X were transferred to a petri dish and dredged with pieces of teased testis for 30 seconds. After leaving for 1 minute a small amount of distilled water was added. Two minutes later the dish was flooded with distilled water.

Fertilised eggs were de-jellied in 2% (w/v) cysteine (adjusted to pH8 with NaOH) and then washed at least 4 times in 1/10x barth-X.

Prior to gastrulation embryos were transferred to 1/10x barth-X to avoid exogastrulation.

Embryos were staged according to Nieuwkoop and Faber (1956).

3.3 Microinjection Of Xenopus Oocytes.

Large oocytes that had been obtained as described in section 3.2 were microinjected using a fine, drawn out glass capillary mounted on a micro manipulator. The needle was linked via a thin gauge pipe to a syringe driven by a vernier, and the system partially filled with oil. Oocytes were transferred to a piece of moist filter paper on a microscope slide and injected with about 10nl of DNA (in sterile water) or RNA (in DEPC treated water) under a binocular microscope. For nuclear injections the needle was inserted a small distance into the animal pole where the germinal vesicle usually lies. Cytoplasmic injections were into the vegetal half of the oocyte. Microinjected oocytes

Methods

were maintained at 20°C in 1x barth-X.

Oocytes injected with synthetic RNA into the cytoplasm were maintained in 1x barth-X at 20°C in petri dishes for a few hours, and then healthy oocytes were transferred into microtitre dishes (so that they could be maintained in a small volume) in 5ul 1x barth-X (containing 5uCi ³⁵S-methionine) per oocyte. Oocytes were maintained overnight in this medium to allow incorporation of ³⁵S-methionine into translation products, and then protein extract (using the method described for band shift assay extract, see section 3.31.1) was prepared from healthy oocytes.

3.4 Preparation of RNA.

3.4.1 From *X. laevis* oocytes.

Oocytes were homogenised in a glass homogeniser in approximately 6 volumes of a 1:1 mixture of NAE (0.3M sodium acetate pH6.5, 0.1M EDTA) containing 2% SDS and neutral phenol. The homogenate was microfuged at 12000 rpm for 5 minutes. The aqueous phase was taken, phenol extracted 2 to 3 more times until it became clear and then chloroform extracted. RNA precipitated with 2 volumes of ethanol at -20°C. RNA was pelleted by microfugation at 12000 rpm for 10 minutes and resuspended in DEPC treated water. RNA concentration was determined by measuring absorbance at 260nm. The RNA was then split into aliquots, re-precipitated with ethanol and stored at -20°C.

Methods

3.4.2 From adult X. laevis liver.

RNA was isolated from adult Xenopus liver using the method of Chirgwin et al (1979). The liver was removed from an animal that had been killed with euthatal and cut into pieces in phosphate buffered saline (PBS). The pieces were washed several times in PBS and then homogenised in five tissue volumes of lysis buffer (6 M guanidinium thiocyanate, 10mM Tris.HCl pH7.5, 1mM B-mercaptoethanol, 4% N-lauroylsarcosine). Caesium chloride was added to a final concentration of 15% (w/v) and the homogenate centrifuged at 2000 rpm in a bench centrifuge to pellet the debris. The supernatant was carefully layered onto 2ml of 5.7M caesium chloride, 100mM EDTA (pH8) in an ultracentrifuge tube. The tube was centrifuged for 12 to 18 hours in a swing out rotor at 36,000 rpm at 20°C. The supernatant was removed with a pasteur pipette until about 1ml remained, this was poured off and the tube inverted to drain. The RNA pellet was resuspended in DEPC treated water and then extracted with an equal volume of neutral phenol. The aqueous phase was precipitated with 2 volumes of ethanol in the presence of 0.3M sodium acetate (pH6.5) at -20°C. The RNA was recovered by centrifugation at 10,000 rpm for 15 minutes and resuspended in DEPC treated water. The RNA concentration was determined by measuring the absorbance at 260 nanometers assuming a reading of 1 for a 40ug/ml solution. Finally, the RNA was ethanol precipitated and stored at -20°C.

Methods

3.4.3 From microinjected oocytes.

Groups of 10 to 20 healthy oocytes were transferred to eppendorf tubes and excess medium removed. 0.5ml of NaE (0.3M sodium acetate pH6.5, 1mM EDTA pH8) was added and the oocytes homogenised by squirting them up and down in a Gilson tip. SDS was then added to a final concentration of 2% (w/v) and the homogenate immediately extracted with an equal volume of neutral phenol. The aqueous phase was taken and extracted with an equal volume of chloroform and then nucleic acids precipitated by adding 2 volumes of ethanol.

3.4.4 Selection Of Polyadenylated RNA By Oligo dT Cellulose Chromatography

This was carried out essentially as described by Maniatis et al (1982). A 2ml oligo (dT) cellulose column was poured in a disposable plastic column and equilibrated with loading buffer (20mM Tris.HCl pH7.6, 0.5M NaCl, 1mM EDTA, 0.1% SDS in DEPC treated water). A column was used several times and stored at 4°C. RNA was dissolved in DEPC treated water and an equal volume of 2x loading buffer (40mM Tris.HCl pH7.6, 1M NaCl, 2mM EDTA, 0.2% SDS) added. This solution was applied to the column, warming the column with a heating lamp as necessary. The column was washed with 5 to 10 column volumes of loading buffer and then with 5 column volumes of wash buffer (20mM Tris.HCl pH7.6, 0.1M NaCl, 1mM EDTA, 0.1% SDS). The flow through was collected in 1 ml aliquots throughout. The A₂₆₀ of these fractions

Methods

was measured as they were collected so that the amount of RNA coming off the column could be assessed. The column was washed with loading buffer and then with wash buffer until the A_{260} was less than 0.1. Polyadenylated RNA annealed to the column was eluted with 2 to 3 column volumes of elution buffer (10mM Tris.HCl pH7.6, 1mM EDTA (pH8), 0.05% SDS). These eluted fractions were pooled and the RNA concentration determined by measuring the A_{260} . Finally the RNA was split into aliquots, precipitated with ethanol and stored at -20°C .

3.5 Gels for Resolving Nucleic Acids.

3.5.1 Non-denaturing Agarose Gels.

DNA samples, to which 0.2 volumes of loading buffer (50% glycerol, 5x TBE, 0.1% bromophenol blue) had been added, were separated in 0.6 to 1.5% (w/v) agarose gels containing 0.2ug/ml ethidium bromide made in 1x TBE buffer. Gels were run in 1x TBE buffer containing 0.5ug/ml ethidium bromide and examined and photographed on an ultraviolet light box.

3.5.2 Low Melting Point Agarose Gels.

These gels were used for the isolation of DNA fragments generated by restriction enzyme digests.

DNA samples to which 0.2 volumes of agarose gel loading buffer had been added were loaded onto a 0.8% (w/v) low melting point agarose gel made in 1x TBE buffer and containing 0.5ug/ml ethidium bromide. Gels were run in 1x

Methods

TBE buffer containing 0.5ug/ml ethidium bromide at a maximum of 40 mA. The gel was examined under U.V. light and the required bands excised as a small gel slice. Exposure of the gel to U.V. light was kept to a minimum to avoid U.V. induced DNA damage. Gel slices were transferred to 1.5ml eppendorf tubes and weighed to calculate the volume of the slice (assuming $1g = 1ml$). 4 volumes of NaE (0.3M sodium acetate pH6.5, 1mM EDTA) was added and the slice melted by heating at 65°C for 10 to 15 minutes. After cooling to 37°C an equal volume of neutral phenol was added, the solution mixed vigorously and left on ice for 10 minutes. The tubes were then centrifuged at 12000 rpm for 5 minutes and the aqueous phase transferred to a fresh tube. The aqueous phase was ether extracted, and the volume reduced to about 300ul by several extractions with butan-1-ol. The DNA precipitated with 2 volumes of ethanol at -20°C.

3.5.3 Formaldehyde Agarose RNA Gels.

RNA samples (up to 20ug) were resuspended in 10ul of DEPC treated water and 20ul of denaturing mix (66% deionised formamide pH7, 1.3x MOPS, 20% formaldehyde) and 3ul of loading mix (50% glycerol, 0.2% bromophenol blue) added. The samples were loaded onto a 1.5% (w/v) agarose, 15% formaldehyde gel made in 1x MOPS buffer (20mM MOPS, 5mM sodium acetate, 1mM EDTA). The samples were electrophoresed at 40mA (100 volts) for 5 to 6 hours in 1x MOPS buffer. The

Methods

gel was stained by soaking in 250ml of 10% glycine for 10 minutes, adding 100ul of 10mg/ml ethidium bromide and soaking for a further 10 minutes. The gel was destained by washing twice, for 15 minutes each, in distilled water and then photographed on an ultraviolet light box.

3.5.4 Denaturing polyacrylamide gels.

6% or 8% polyacrylamide (19:1 bis) gels containing 42% (w/v) urea in 1x TBE were poured between 20x 40cm gel plates with 4mm spacers. Nucleic acid samples in denaturing gel loading buffer (90% deionised formamide, 10mM EDTA pH8, 0.01% bromophenol blue, 0.01% xylene cyanol) were heated at 100°C for 5 minutes, loaded onto the gel and electrophoresed at 38 watts in 1x TBE.

Gels with ³⁵S samples (eg. sequencing reactions) were fixed in 10% ethanol (v/v), 10% acetic acid (v/v) for 15 minutes, transferred to a sheet of blotting paper, dried at 80°C on a vacuum drier and then exposed to X-ray film at room temperature.

Gels with ³²P samples (eg. primer extension analysis, RNase protection assays) were generally not fixed and dried before autoradiography. The top gel plate was removed, the gel covered in cling film and exposed to X-ray film with an intensifying screen at -70°C.

Methods

3.6 Isolation Of Genomic DNA From Adult Xenopus Blood.

An adult female was anaesthetised by immersion in 0.2% (w/v) MS222 for 20 to 30 minutes. The thorax was then opened on the ventral side to expose the heart. A 10 ml syringe was filled with 3ml of 1mg/ml heparin in NMT (100mM NaCl, 10mM Tris.HCl pH7.4, 3mM MgCl₂) and the needle inserted into the still beating heart. Blood was gently collected into the syringe and transferred to cold universals which were centrifuged at 2000 rpm for 5 minutes. The cell pellet was resuspended in 10ml of cold NMT and re-centrifuged. This was repeated until the supernatant was no longer pink. The cells were lysed by resuspending in 10ml of distilled water and leaving at room temperature for 5 minutes. 100ml of neutral phenol was transferred to a glass bottle and the lysed cell suspension added to this. After mixing vigorously the DNA was sheared by twice taking the emulsion into a syringe through a coarse grade needle. The emulsion was transferred to oakridge tubes and centrifuged at 8000 rpm for 20 minutes at room temperature. The aqueous phase was carefully transferred to a fresh tube and re-extracted with an equal volume of neutral phenol. The aqueous phase was transferred to a glass beaker and approximately two volumes of ethanol was added to precipitate the DNA. High molecular weight DNA was spooled onto the end of a pasteur pipette and transferred to a plastic universal. The DNA was dissolved in 10ml of TE (10mM Tris.HCl pH7.5, 1mM EDTA) by shaking

Methods

gently at room temperature for several hours. Ribonuclease A was added to a final concentration of 5ug/ml and the solution incubated at 37°C for 30 minutes. An equal volume of 2x PK buffer (0.2 M Tris.HCl pH7.5, 25mM EDTA, 0.3M NaCl, 2% (w/v) SDS) and 0.1 volumes of 2mg/ml proteinase K were added and the solution incubated at 37°C for 1 hour. The solution was extracted with an equal volume of neutral phenol and the phases separated by centrifugation at 8000 rpm for 20 minutes. The aqueous phase was transferred to corex tubes and the DNA precipitated with ethanol and recovered by centrifuging at 10 000 rpm for 20 minutes. The pelleted DNA was resuspended in 5ml of TE and the concentration determined by measuring the absorbance at 260 nanometers, assuming that a 50ug/ml solution has an A₂₆₀ of 1. The DNA solution stored at 4°C.

3.7 Southern Blotting.

DNA samples separated in non-denaturing agarose gels as described in section 3.5.1 were transferred to nitrocellulose using the method of Southern (1975). To increase the efficiency of transfer of high molecular weight DNA (especially with genomic Southern blots) the gel was first treated with 0.25M HCl for 15 minutes at room temperature. This partially hydrolyses the DNA, generating smaller fragments which transfer more efficiently. After rinsing with distilled water the gel was incubated in a large volume of denaturing solution (1.5M NaCl, 0.5M NaOH)

Methods

with gentle shaking for 20 minutes. This was repeated with fresh denaturing solution. The gel was rinsed with distilled water and then neutralised by incubating with gentle shaking in two changes (for 30 minutes each) of a large volume of neutralising solution (3M NaCl, 0.5M Tris.HCl, 1mM EDTA, pH7). A piece of filter paper (Whatman 3MM) moistened with 20x SSC was placed on top of a platform in a tray of 20x SSC so that the ends of the filter paper were submerged in the 20x SSC. The gel was placed on top of the filter paper and the exposed filter paper covered with cling film. A piece of nitrocellulose was cut to the same size as the gel, moistened with 2x SSC and lowered onto the surface of the gel, making sure that no air bubbles formed. Two pieces of filter paper cut to the same size as the gel were soaked in 2x SSC and lowered onto the surface of the nitrocellulose. Tissue paper was stacked on top of this and weighed down with a brick. The gel was left to blot overnight after which the nitrocellulose filter was rinsed briefly in 2x SSC. DNA was fixed to nitrocellulose filters by baking at 80°C under vacuum for 2 hours.

Prehybridisation and hybridisation was carried out in heat sealed plastic bags weighed down in a water bath set to the appropriate temperature. Filters were prehybridised in 20 to 50ml (depending on the size of filter) of 5x Denhardt's, 6x SSC, 0.1% SDS, 100ug/ml E. coli tRNA, 50% deionised formamide at 42°C for 3 hours to overnight. The prehybridisation solution was discarded and replaced with

Methods

10 to 20ml of the same buffer containing radiolabelled nucleic acid probe. The bag was re-sealed and hybridised at 42°C overnight.

Filters were first washed twice in 2x SSC, 0.1% SDS at room temperature for 5 minutes each and then twice for 30 minutes each in the final washing conditions (which depended upon the hybridisation stringency required). Finally, the filters were blotted on filter paper and while still damp, wrapped in cling film and exposed to X-ray film with an intensifying screen at -70°C.

3.8 Northern Blotting.

After staining with ethidium bromide formaldehyde agarose RNA gels were soaked in 250ml of 20x SSC for 20 minutes and the RNA transferred to nitrocellulose (Hybond-C, Amersham) as described by Thomas (1980). The gel was capillary blotted overnight using 20x SSC as the transfer buffer exactly as described for Southern blotting (see section 3.7) after which the nitrocellulose filter was briefly dried at room temperature and then baked at 80°C under vacuum for 2 hours.

Prehybridisation, hybridisation, washing and autoradiography were exactly as described for Southern blotting (see section 3.7).

Methods

3.9 Large Scale Preparation Of Plasmid DNA And Purification By Caesium Chloride/Ethidium Bromide Centrifugation.

The method used was the alkaline lysis method as described by Maniatis et al (1982). A single colony was inoculated into 10ml of sterile LB containing the appropriate antibiotic, and grown overnight at 37°C. The next day 0.5ml of this overnight culture was used to inoculate 50ml of LB, and this culture grown on a shaker at 37°C for 2 to 3 hours. This 50ml culture was used to inoculate 500ml of LB in a 2 litre flask and this culture was shaken (200 rpm) overnight at 37°C. The next day, the cells were pelleted by centrifugation at 5000 rpm for 10 minutes. The following quantities are those used for cells from a single 500ml bacterial culture. The pellets were resuspended in a total of 10ml of ice cold solution I (50mM glucose, 25mM Tris.HCl (pH8), 10mM EDTA (pH8), 5mg/ml lysozyme) and equal volumes transferred to two oakridge centrifuge tubes. After incubating at room temperature for five minutes 10ml of freshly prepared solution II (0.2M NaOH, 1% SDS) was added to each tube and the contents mixed by inverting several times. The tubes were left on ice for 10 minutes after which 7.5ml of ice cold 5M potassium acetate (pH6) was added to each tube. The tubes were mixed by inverting several times and left on ice for 10 minutes. The pH of the solution was then tested with pH paper to ensure that the sodium hydroxide had been completely neutralized, and more potassium acetate added if necessary. Bacterial debris and

Methods

chromosomal DNA were pelleted by centrifugation at 13 000 rpm for 30 minutes. 18ml of supernatant from each tube was transferred to a 30ml corex tube and nucleic acids precipitated by adding 12ml of isopropanol. After 15 minutes at room temperature the tubes were centrifuged at 10 000 rpm for 30 minutes at 20°C, and the pellets resuspended in a total of 20ml of TE. Exactly 20g of caesium chloride was dissolved in this and 0.6ml of ethidium bromide (10mg/ml) added. Using a syringe, this solution was transferred to a Beckman heat-sealable centrifuge tube. The tubes were topped-up with liquid paraffin, balanced to within 10mg and heat sealed. They were then centrifuged at 45 000 rpm in a vertical rotor for 18 hours at 20°C.

The tubes were viewed under U.V. light and the lower band (which is supercoiled plasmid DNA) was removed from the gradient using a syringe. The solution was extracted at least three times with water saturated butan-1-ol to remove the ethidium bromide, and then dialysed against 2 litres of TE at 4°C overnight. The DNA was precipitated in corex tubes at -20°C by the addition of sodium acetate to a final concentration of 0.3M and 2 volumes of ethanol. The DNA was recovered by centrifugation at 10,000 rpm for 30 minutes at 4°C and resuspended in 0.5ml of TE. The solution was then extracted with neutral phenol and re-precipitated with ethanol. The DNA was pelleted by microfugation for 10

Methods

minutes and dissolved in 0.5ml of TE. The DNA concentration was determined by measuring the A₂₆₀.

3.10 Primer Extension Analysis of RNA.

The single stranded oligonucleotide to be used as the primer was end labelled with γ -³²P-ATP (see section 3.29.2) and resuspended in 3x PEB (1x PEB is 0.4M NaCl, 10mM PIPES (pH6.4), 0.5mM EDTA (pH8)). Hybridisation reactions were set up by mixing the RNA sample (in 20ul DEPC treated water) and 10ul of the labelled primer (in 3x PEB) in a 0.5ml eppendorf tube. The tubes were tightly sealed and incubated overnight at the hybridisation temperature. For 17mer and 18mer primers 45°C was found to be optimal.

After hybridisation 60ul of ethanol was added and the tubes were left at -20°C for 30 minutes. Nucleic acids were recovered by microfugation for 10 minutes and dissolved in 20ul of reverse transcriptase reaction buffer (50mM Tris.HCl pH8.3, 6mM MgCl₂, 10mM DTT, 1mM each of dATP, dCTP, dTTP, dGTP). 9 units of AMV reverse transcriptase was added and the reactions incubated at 37°C for 30 minutes. Ribonuclease A was added to a final concentration of 2ug/ul and the tubes incubated for a further 10 minutes at 37°C. DNA was precipitated at -20°C by adding sodium acetate (pH6.5) to a final concentration of 0.3M and 2 volumes of ethanol. After microfugation the pellet was resuspended in 4ul of denaturing gel loading dye (90% deionised formamide, 10mM EDTA (pH8), 0.01% xylene cyanol, 0.01% bromophenol

Methods

blue), heated at 100°C for 5 minutes and then loaded on to a 8% denaturing polyacrylamide (see section 3.5.4). The gel was run at 38 watts until the bromophenol blue was about 2 inches from the bottom of the plate. The plates were prised apart, the gel covered with cling film and exposed to x-ray film with an intensifying screen at -70°C.

3.11 RNase Protection Assays.

RNase protection assays were carried out essentially as described by Zinn et al (1983). 2.5×10^5 cpm of internally labelled RNA probe (see section 3.29.4) was hybridised to RNA in 30ul of S1 hybridisation buffer (40mM PIPES pH6.4, 1mM EDTA pH8, 0.4M NaCl, 80% deionised formamide) at 45°C overnight.

Next day 350ul of digestion buffer (10mM Tris-HCl pH7.5, 5mM EDTA, 300mM NaCl) containing RNases was added and the reaction incubated at 37°C for 1 hour. A range of RNase concentrations were tried. 20ug/ml RNaseA (Sigma) and 400 units/ml RNaseT1 (Boehringer) seemed to give the best results. 50ug of proteinase K (Boehringer) and 20ul of 20% SDS were added and incubation continued for a further 15 minutes. The reactions were extracted with neutral phenol and precipitated with 2 volumes of ethanol and carrier tRNA (10ug). After microfugation samples were suspended in denaturing gel loading buffer, boiled for 5 minutes and electrophoresed on a denaturing polyacrylamide gel (see

Methods

section 3.5.4) at 38 watts until the bromophenol blue was about 1cm from the bottom of the gel.

3.12 Subcloning Techniques.

3.12.1 Restriction enzyme digests.

These were carried out according to the manufacturer's instructions. Plasmid DNAs were generally digested for 1 hour, whilst genomic DNA for southern blotting was digested for 5 hours, adding a second aliquot of enzyme after part way through the incubation.

3.12.2 Preparation of plasmid vectors for sub-cloning.

Vectors that had been digested with two enzymes generating incompatible ends were run on a low melting point agarose gel and the linear vector fragment recovered. Vectors cut with a single restriction enzyme were treated with calf intestinal alkaline phosphatase (CIAP) by adding 1 or 2 units of enzyme into the digestion reaction at the end of the digestion period and incubating for a further 30 minutes at 37°C. 200ul of NAE was added, followed by extraction with neutral phenol, extraction with ether and precipitation with 2 volumes of ethanol.

Methods

3.12.3 Preparation of target DNA for subcloning.

Target DNA was generally a restriction fragment isolated from a low melting agarose gel (see section 3.5.2).

For some purposes it was necessary to ligate target and vector with incompatible sticky ends. In this case 5' overhangs (on both target and vector) were 'blunted' by the addition of all 4 dNTPs (final concentration 1mM) and 10 units of Klenow fragment of DNA polymerase into the digestion reaction (at the end of the digestion period) and incubating for a further 30 minutes at room temperature.

3.12.4 Ligations.

10ul reactions were carried out in 1x C buffer (66mM Tris.HCl pH7.6, 6.6mM MgCl₂, 10mM DTT), containing 1mM ATP, vector DNA, target DNA and T4 DNA ligase. Usually, 20ng of vector and a concentration range of target DNA was used. A control ligation that contained vector alone was always included. 0.1 units of T4 DNA ligase was used for sticky end ligations and 1 unit for blunt end ligations. The reactions were incubated at 14°C overnight.

3.13 Transformation of E.coli with Plasmids.

3.13.1 CaCl₂ Mediated Transformation.

This method was used for the E. coli strains MC1061, TG2, 71:18 mutL and BL21 (DE3). An appropriate volume of 2x TY or LB was inoculated with 0.01 volume of an overnight culture of the host bacteria and incubated at 37°C with

Methods

vigorous shaking until the A_{550} reached 0.5. The culture was then cooled on ice for 10 minutes and the cells pelleted by centrifugation at 2000 rpm for 5 minutes. The pellet was resuspended in 0.5 volume of ice cold 50mM CaCl_2 and left on ice for 20 minutes. The cells were re-pelleted and resuspended in 0.1 volume of ice cold 50mM CaCl_2 . After incubating the cell suspension on ice for a further 10 minutes (or longer) 0.2ml aliquots were transferred to eppendorf tubes on ice. Half of the appropriate ligation reaction (5ul) or approximately 1ng of uncut plasmid DNA was added to an aliquot, followed by mixing and incubation on ice for 40 minutes. The cells were then heat shocked at 42°C for 2 minutes, transferred to ice for 5 minutes, and spread onto dried L-agar plates (containing the appropriate antibiotic) with an ethanol sterilized glass spreader. The plates were incubated inverted at 37°C overnight.

3.13.2 Hanahan High Efficiency Transformation.

This method was used for the transformation of JM101 with plasmid and M13 vectors and is essentially as described by Hanahan (1983). An appropriately sized culture of host cells was grown as described in section 3.13.1, and when the A_{550} reached between 0.3 and 0.4 the culture was cooled on ice for 15 minutes. The cells were then pelleted by centrifugation at 2000 rpm for 5 minutes and resuspended in 0.3 original culture volume of transformation buffer (TFB; 10mM KHES pH6.3, 100mM KCl, 45mM $\text{MnCl}_4\cdot\text{H}_2\text{O}$, 10mM CaCl_2 , 3mM

Methods

HACoCl₃). The suspension was left on ice for 15 minutes after which the cells were pelleted and resuspended in 0.08 original culture volume of TFB. At this point the cell suspension was transferred to a polypropylene tube and left on ice for 10 minutes. Dimethylformamide (DMF) was added to a final concentration of 4.75% (v/v), the suspension mixed well and left on ice for 10 minutes. A solution of 2.25M DTT in 10mM potassium acetate was added to 4.1% (v/v), the suspension mixed well and left on ice for a further 10 minutes. A second aliquot of DMF was then added so that the final concentration of DMF was 9.5% (v/v). The suspension was incubated on ice for 20 minutes and then 0.2ml aliquots transferred to eppendorf tubes on ice. Half of the appropriate ligation reaction (5ul) or approximately 1ng of uncut plasmid DNA was added to an aliquot, followed by mixing and incubation on ice for 40 minutes. The cells were then heat shocked at 42°C for 2 minutes, transferred to ice for 5 minutes. The transformed cells were then spread onto dried L-agar plates (containing the appropriate antibiotic) with an ethanol sterilized glass spreader, or in the case of M13 transformations plated in top agarose (see section 3.14). The plates were incubated inverted at 37°C overnight.

Methods

3.14 Plating E.coli Transformed Or Infected With Bacteriophage M13.

For each transformation or infection, 3ml of molten H-top agar was pipetted into a sterile glass tube in a heating block at 45°C. After allowing the top agar to equilibrate to 45°C 30ul of X-gal (20mg/ml in DMF), 30ul of IPTG (24mg/ml in water) and 200ul of exponential host cells were added, in that order. The transformed or infected host cells were then added, the contents of each tube mixed by inversion and poured onto a dried H-agar plate. The plate was rocked gently to ensure an even covering of top agar and then allowed to set at room temperature for 10 minutes. Plates were incubated inverted at 37°C overnight.

3.15 Small Scale Isolation Of Plasmid DNA And M13 RF DNA.

Single bacterial colonies were inoculated into 2ml of 2x TY containing the appropriate antibiotic using a sterile toothpick, and grown in a shaking incubator at 37°C for 8 to 14 hours. For preparing M13 RF DNA 2ml of 2x TY was inoculated with 20ul of an overnight culture of the host cells and 10ul of phage stock or a sterile toothpick which had been stabbed into a single plaque. M13 cultures were grown at 37°C with vigorous shaking (300 rpm) for 6 to 8 hours. The cultures were then transferred to 1.5ml eppendorf tubes and then microfuged for 5 minutes. In the case of M13 infected cells an aliquot of the supernatant

Methods

(which contains bacteriophage particles) was removed and stored at -20°C . This supernatant was also used for the preparation of single stranded M13 DNA (see section 3.16). The pellets were resuspended in 0.2ml of STET (8% sucrose (w/v), 0.5% triton x-100 (v/v), 50mM EDTA (pH8), 10mM Tris.HCl (pH8)) and 10ul of fresh 10mg/ml lysozyme added. The tubes were incubated on ice for 10 minutes, and then heated at 100°C for 40 seconds. Bacterial debris and chromosomal DNA were pelleted by microfugation for 10 minutes. This viscous pellet was removed using a toothpick and nucleic acids in the supernatant precipitated by adding 0.1 volumes of 3M NaAc (pH6.5) and 1 volume of isopropanol. The tubes were left at -20°C for at least 30 minutes and then microfuged for 10 minutes to recover the nucleic acids. The pellet was resuspended in 0.2ml of NaE (0.3M sodium acetate pH6.5, 1mM EDTA) and the solution extracted with an equal volume of neutral phenol. The aqueous phase was extracted with ether and nucleic acids precipitated by adding 2 volumes of ethanol. After at least 30 minutes at -20°C nucleic acids were recovered by microfugation for 10 minutes and the pellet resuspended in 40ul of TE. If necessary, RNA was removed by treatment with RNase A at a final concentration 20ug/ml for 15 minutes at 37°C .

Methods

3.16 Preparation Of Single Stranded M13 Template DNA.

2ml of 2x TY was inoculated with 20ul of an overnight culture of the host strain, usually JM101. Phage particles from a single plaque were transferred to the media using a sterile toothpick, or alternatively 10ul of a frozen phage stock was added. Cultures were then incubated at 37°C with vigorous shaking (300 rpm) for 6 to 8 hours. The bacteria were pelleted by microfugation for 5 minutes and 1ml of the supernatant transferred to an eppendorf tube. Bacteriophage particles were then precipitated by adding 200ul of 20% (w/v) polyethylene glycol in 2.5 M NaCl and leaving the tubes at room temperature for 20 minutes. The phage were pelleted by microfugation for 10 minutes and the supernatant removed using a Gilson pipette. The tubes were re-microfuged for 10 seconds to bring down the remaining liquid, which was carefully removed using a drawn out glass capillary. The pellet was resuspended in 0.2ml of NaE (0.3M sodium acetate pH6.5, 1mM EDTA) and the suspension extracted with an equal volume of neutral phenol. The aqueous phase was extracted with an equal volume of chloroform and the DNA precipitated by adding 2 volumes of ethanol. After at least 30 minutes at -20°C the DNA was recovered by microfugation for 10 minutes and dissolved in 10ul of TE. Template DNA was checked (for concentration and contamination with RNA) on a 1% non-denaturing agarose gel prior to sequencing.

Methods

3.17 Complementation Tests On Single Stranded M13 DNA

For M13 subclones (intended for sequencing) in which the cloned fragment could have been inserted in either orientation, and it was necessary to sequence from each end of the cloned fragment, complementation tests on single stranded phage DNA were used to select clones of both orientations.

10ul of phage supernatant (in 2x TY) from each of two clones was mixed with 2ul of 1% SDS and hybridised at 65°C for 1 hour. After adding 4ul of loading buffer the samples were analysed on a 1% non-denaturing agarose gel. The presence of a slower migrating band (not present in either phage supernatant alone) indicated that the two clones had annealed, and therefore contained the insert in opposite orientation.

3.18 Plating Bacteriophage Lambda.

An appropriate volume of 2x TY was inoculated with host cells (BB4 or Y1090) from an overnight culture and grown on a shaker at 37°C until the A_{550} reached 0.5. The cells were pelleted by centrifugation at 2000 rpm for 5 minutes and resuspended in 0.05 original culture volume of 10mM $MgSO_4$. Bacteriophage particles (from phage stock in SM buffer (50mM Tris, HCl pH 7.5, 0.58% (w/v) NaCl, 0.2% (w/v) $MgSO_4 \cdot 7H_2O$, 0.01% (w/v) gelatin)) were mixed with these host cells and allowed to adsorb for 15 minutes at 37°C. Up to 100 000 phage were absorbed with 0.6ml of host cells. 6ml of molten

Methods

top agarose was aliquoted in to sterile tubes in a heating block at 45°C and allowed to equilibrate to 45°C. 0.6ml of infected cells was added to the top agarose and this was then poured onto dried 10cm x 10cm NZY plates and allowed to set on a level surface. The plates were then incubated inverted at 37°C until plaques of a suitable size formed.

3.19 Bacteriophage lambda Plaque Lifts.

This procedure was used to transfer bacteriophage DNA from plates to nitrocellulose filters and is essentially as described by Benton and Davis (1975). The host strains BB4 or Y1090 were used. Bacteriophage lambda were plated (see section 3.18) using top agarose (not agar), and then incubated at 37°C for 5 to 6 hours until suitably sized plaques formed. Plaques became visible about 4 hours after plating and were allowed to grow for a further 1 to 2 hours. The plates were then transferred to 4°C for at least 2 hours. Pieces of nitrocellulose, cut slightly smaller than the plates, were carefully lowered onto the surface of the top agarose ensuring that no air bubbles formed. These were left in place for 1 minute during which time the position of the filter was marked by making three asymmetrical holes through the filter and into the agar with a syringe needle. The position of the holes was also marked on the base of the plates using a marker pen. The filters were carefully removed with a pair of forceps and placed, DNA side up, onto a tray containing several sheets

Methods

of filter paper (Whatman 3MM) moistened with denaturing solution (0.5M NaOH, 1.5M NaCl). The filters were left in place for 1 minute during which time a second nitrocellulose filter was applied to each plate. This replica filter was left in place for 2 minutes and orientated by making holes in the identical positions used for the first filter. The replica filter was denatured in the same way as the first filter. After denaturation the filters were transferred to a tray containing several sheets of 3MM moistened with neutralising solution (3M NaCl, 0.5M Tris.HCl pH7, 1mM EDTA), left in place for 4 minutes and then transferred to a second tray of neutralising solution for a further 5 minutes. Finally filters were transferred onto a tray of 3MM soaked in 2x SSC for 5 to 10 minutes. The filters were then allowed to air dry.

The filters were then baked at 80°C under vacuum for 2 hours. Prehybridisation, hybridisation and autoradiography of filters was as described for Southern blotting (see section 3.7)

The region containing a positive plaque was identified by the orientation marks made in the filter and plate. Agar plugs were cut from the plate, transferred to 1ml of SM buffer containing 3ul of chloroform and stored at 4°C. This phage stock was replated (at a lower density) and plaque lifts performed. Positive plaques were identified and the

Methods

process repeated until the agar plug taken contained a single positive plaque.

3.20 Automatic Excision Of Phagemids From lambda ZAP Clones.

This was carried out exactly as described by the manufacturer. A culture of BB4 was grown until the A_{550} was between 0.8 and 1. 0.2ml of culture was mixed with 100ul of lambda ZAP phage stock (from a single positive plaque, see section 3.19) and 10ul of R408 helper phage (7×10^{10} pfu/ml, supplied by Stratagene). The cultures were incubated in a 37°C shaker at 200 rpm for 4 to 6 hours, heated at 70°C for 20 minutes and then centrifuged at 2,000 rpm for 5 minutes. The supernatant, which contains the phagemid particles, was transferred to a fresh tube and stored at 4°C. The phagemid was converted to colony form by mixing 0.2ml of a fresh BB4 culture (A_{550} about 0.5) with 0.1 to 100ul of the phagemid stock, incubating at 37°C for 15 minutes to allow adsorption of the phage, and then spreading 10 and 100ul aliquots onto dried L agar plates containing ampicillin (100ug/ml). The plates were incubated inverted overnight at 37°C.

Methods

3.21 DNA Sequencing by the Dideoxy Chain Termination Method.

3.21.1 Sequencing single stranded M13 templates.

This method was first described by Sanger et al (1977).

A set of deoxynucleotide and dideoxynucleotide mixes were assembled as shown below:

(A)	A ^o	C ^o	G ^o	To
0.5mM dCTP	20ul	1ul	20ul	20ul
0.5mM dGTP	20ul	20ul	1ul	20ul
0.5mM dTTP	20ul	20ul	20ul	1ul
TE buffer	20ul	20ul	20ul	20ul

(B) Dideoxy NTP working solutions: 0.03mM ddATP, 0.1mM ddCTP, 0.075mM ddGTP, 0.5mM ddTTP. These concentrations were altered as necessary to suit a particular template, or reduced (generally by half) to allow reading of sequence at a greater distance from the primer.

(C) Chase mix: 0.5mM of all 4 dNTPs.

For each clone to be sequenced an annealing reaction was set up by mixing 1ul of single stranded template DNA, 6.5ul of sterile water, 1.5ul of freshly prepared klenow reaction buffer (9ul of 10x core buffer (100mM Tris.HCl (pH8), 50mM MgCl₂) and 1ul of 700mM B-mercaptoethanol), and 1ul of M13 17mer sequencing primer (2ng/ul). To anneal the primer to template the tube was placed in a large shallow tray

Methods

containing water at 80 to 85°C and left to cool at room temperature for about 30 minutes. For each template, four 0.5ml eppendorf tubes were labelled A, C, G, and T. 1ul of the A⁰ nucleotide mix and 1ul of the ddATP nucleotide solution was placed into the "A" tube. This was repeated for the "C", "G", and "T" tubes using the appropriate nucleotide mixes. The annealed template/primer mix was centrifuged briefly to bring down any condensation, and then 1ul of α -³⁵S-dATP and 1ul of DNA polymerase Klenow fragment (5 units/ul) added. 2.5ul of this mix was added to each of the tubes containing the nucleotide mixes. The reaction was allowed to proceed at 30°C for 20 minutes after which 1ul of chase mix was added and the tubes incubated for a further 10 minutes at 30°C. The reaction was terminated by adding 5ul of denaturing gel loading buffer.

One third of each reaction was analysed on a 6% denaturing polyacrylamide gel (see section 3.5.4) made with a sharks tooth comb. The gel was generally run until the bromophenol blue in the loading buffer was at the bottom, and then a second 1/3 of the reaction loaded. Electrophoresis was continued until the bromophenol blue from this loading was at the bottom of the gel, and if necessary the process repeated with the final 1/3 of the reaction.

3.21.2 Sequencing plasmid DNA.

Plasmid DNA from a small scale isolation treated with RNaseA (see section 3.15) was used as template.

Methods

Approximately 2ug of DNA in 200ul of NAE was extracted with neutral phenol and precipitated with 2 volumes of ethanol at -20°C . DNA was recovered by microfugation for 10 minutes and suspended in 20ul of 0.2M sodium hydroxide; 0.2mM EDTA. The solution was left at room temperature for 5 minutes, and then DNA precipitated by adding 2ul of 2M ammonium acetate (pH4.5) and 100ul of ethanol and leaving at -70°C for 30 minutes. DNA was recovered by microfugation for 15 minutes, and then the pellet was washed in 80% (v/v) ethanol and briefly dried under vacuum. DNA was suspended in 10ul of a solution made by mixing 1.5ul of reaction buffer (see section 3.21.1), 6.5 ul of water and 2ul of M13 17mer sequencing primer (2ng/ul). Primer was annealed to template by incubation at 37°C for 15 minutes, and then sequencing reactions carried out as described for single stranded M13 templates (see section 3.21.1).

3.22 Oligonucleotide directed mutagenesis.

This procedure (first described by Kunkel (1985)) was used to introduce point mutations into M13 subclones.

3.22.1 Phosphorylating synthetic oligonucleotides.

Synthetic oligonucleotides were supplied dephosphorylated at the 5' end, and to allow ligation at the end of the extension reaction had to be kinased. 120ng of oligonucleotide was kinased in 20ul of 1x C buffer (66mM Tris.HCl pH7.6, 6.6mM MgCl_2 , 10mM DTT) containing 2mM ATP

Methods

and 20 units of T4 polynucleotide kinase. The reaction was allowed to proceed at 37°C for 1 hour.

3.22.2 Annealing oligonucleotide to template.

6ul of single stranded M13 template (see section 3.16), 1.5ul of 10x TM buffer (100mM Tris.HCl pH7.5, 50mM MgCl₂), 6.5ul of water and 3ul of the oligonucleotide kinasing reaction were mixed, heated to 65°C for 10 minutes and then cooled slowly (over 30 minutes) to room temperature. An equivalent reaction with no oligonucleotide was also set up.

3.22.3 Extension reactions.

3ul of extension mixture (3ul 10x TM, 3ul 5mM all 4 dNTPs, 2ul 10mM ATP, 3ul 100mM DTT, 7ul water) and 5ul of annealing reaction were mixed together. 5 units of Klenow fragment of DNA polymerase and 1 unit of T4 DNA ligase were added, and the reaction incubated at 14°C for 5 hours to overnight. A control reaction with no T4 DNA ligase was also set up.

These reactions, and some template DNA, were electrophoresed on a low melting point agarose gel (see section 3.5.2) and viewed under UV light. The control reaction from an 'annealing' with no oligonucleotide should give a single band of the same mobility as template DNA. The control reaction with no ligase added generally gave a high molecular weight band (assumed to consist of multimers

Methods

generated by strand displacement) and a second, faster migrating band above the position of the template band (assumed to consist of nicked, circular M13 duplexes). The reaction with ligase added generally had an additional band (above the nicked, circular M13 duplexes and below the multimers) which was assumed to consist of closed, circular M13 duplex. This band was excised and the DNA purified.

3.22.4 Screening for mutants.

Purified closed, circular duplex M13 was transformed into E. coli 71:18 mutL as described in section 3.13.1. Single plaques were picked, grown-up and phage stocks taken (see section 3.15).

These phage stocks (which contain single stranded DNA in phage particles) were screened for mutants. 1ul of these phage stocks (and phage stock of the original M13 subclone) were spotted onto duplicate nitrocellulose filters. These filters were then processed as described for bacteriophage lambda plaque lifts (see section 3.19), treating them with denaturing solution and neutralising solution and then baking at 80°C under vacuum for 2 hours.

Filters were then prehybridised and hybridised to end-labelled oligonucleotide (see section 3.29.2 for method of end-labelling oligonucleotides) as described for Southern blotting (section 3.7), except that no formamide was added to hybridisation/pre-hybridisation solutions and hybridisation was carried out at room temperature.

Methods

The duplicate filters were washed at different temperatures in 6x SSC containing 0.1% (w/v) SDS. The melting temperature of the oligo:mutant-template duplex in °C was assumed to be $4(G, C \text{ content of oligo}) + 2(A, T \text{ content of oligo})$. One filter was washed at T_m-2 , and one filter at T_m-15 for 30 minutes to 1 hour. A hybridisation signal should be seen with mutant phage after washing at both T_m-2 and T_m-15 . A hybridisation signal should be seen with non-mutant phage (and the original subclone) only at T_m-15 , this signal being washed off at T_m-2 (since the oligo is mis-matched). The yield of mutants ranged between 30% and 100% of phage stocks tested.

Some phage stocks were found to consist of a mixture of wild-type and mutant phage. Consequently phage were purified by infecting E. coli JM101 with the phage stock, plating-out (see section 3.14), taking single plaques from these plates and re-screening.

Finally the mutants were checked by sequencing single stranded M13 template (prepared as described in section 3.16), or if the mutation generated a restriction endonuclease site by preparing phage RF DNA (see section 3.15) and checking for the presence of that site.

Methods

3.23 SDS-Protein Polyacrylamide Gel Electrophoresis.

3.23.1 Running protein gels.

A 20cm x 20cm denaturing gel was poured in two stages. Firstly the 16cm separating gel, which consisted of 10 to 15% acrylamide (37.5:1 bis), 0.3M Tris.HCl pH8.8, 0.08% (w/v) SDS. After this had set it was overlaid with the 4cm stacking gel (into which the slot former was inserted), which consisted of 3% acrylamide (20:1 bis), 0.125M Tris.HCl pH6.8, 0.1% (w/v) SDS. Protein samples were mixed with an equal volume of 2x loading buffer (0.125M Tris.HCl pH6.8, 20% glycerol, 4% (w/v) SDS, 10% (v/v) 2-mercaptoethanol, 0.002% bromophenol blue), heated at 90°C for 5 minutes and loaded onto the gel. The gel was run at 10 to 40mA in 1x running buffer (50mM Tris, 400mM glycine, 0.1% (w/v) SDS).

Gels were then transferred to nitrocellulose (see section 3.24), or stained for protein with Coomassie blue, or fluorographed and then autoradiographed.

3.23.2 Staining protein gels with Coomassie blue.

Gels were incubated in stain (10% (v/v) acetic acid, 45% (v/v) methanol, 0.1% (w/v) Coomassie blue) for at least 1 hour on a slow shaker. The gel was then incubated in several changes of destain (10% (v/v) acetic acid, 45% (v/v) methanol) on a slow shaker, until background staining was adequately reduced. Gels were photographed wet, and then dried on a vacuum drier at 60°C.

Methods

3.23.3 Fluorography of protein gels.

Gels used to resolve proteins labelled with ^{35}S -methionine were fluorographed to enhance the signal seen on X-ray film.

Gels were incubated in fixer (10% (v/v) acetic acid, 45% (v/v) methanol) for at least 1 hour and then transferred to En^3Hance (Dupont) for 30 to 45 minutes. The gel was then incubated in water (to precipitate the scintillant) for 30 to 45 minutes, dried on a vacuum drier at 60°C , and exposed to X-ray film at -70°C .

3.24 Western Blots.

Proteins resolved on SDS polyacrylamide gels (see section 3.23.1) were electrophoretically transferred to nitrocellulose. The gel was soaked in blotting buffer (20mM Tris, 150mM glycine, 20% (v/v) methanol) for 30 minutes. A sandwich of 3 sheets of filter paper (Whatman 3MM) followed by a sheet of nitrocellulose, the gel and 3 sheets of 3MM was assembled. All components were pre-wetted in blotting buffer. The sandwich was submerged in blotting buffer, in an electro-blotting tank, with the nitrocellulose between the gel and the positive electrode. The gel was blotted at 30 volts overnight (or 60 volts for 3 hours). Blots were stored wrapped in cling film at -20°C .

Methods

3.25 Immuno-detection of Proteins on Western Blots.

A Western blot (see section 3.24) was blocked by incubating in 100ml of TBS (20mM Tris.HCl pH7.6, 137mM sodium chloride), containing 0.1% tween-20 and 5% dried non-fat milk, on a slow shaker for 1 hour. The filter was washed 3 times in 100ml of fresh TBS, containing 0.1% tween-20, for 5 minutes each on a slow shaker. The filter was then incubated in a plastic bag with 10ml of TBS containing antiserum or affinity purified antiserum (usually 100ul) and incubated on a slow shaker for 1 to 2 hours. The filter was then washed, as before, in TBS containing 0.1% tween-20.

Bound antibodies were detected using an Amersham Blotting Detection Kit (for rabbit antibodies) exactly according to the manufacturer's instructions. The filter was incubated in a plastic bag with 10ml of TBS containing biotinylated anti rabbit immunoglobulin on a slow shaker for 30 minutes. The filter was then washed as before. The filter was incubated in a plastic bag with 10ml of TBS containing streptavidin-alkaline phosphatase conjugate on a slow shaker for 20 minutes. The filter was then washed as before. Finally, the filter was incubated in a plastic bag with 10ml of diethanolamine buffer (100mM diethanolamine.HCl pH9.5, 5mM MgCl₂) containing nitro-blue tetrazolium (NBT) and 4-chloro-3-indoyl phosphate (BCIP), developed as long as necessary, and then the reaction stopped by washing in water.

Methods

3.26 Generation and Purification of a Polyclonal Antiserum.

3.26.1 Purification of a fusion protein produced in E. coli for use as an antigen.

A fusion protein produced from a cloned fragment of X. laevis Oct-1 in the pET-3B vector (Rosenberg et al (1987)) was purified for use as an antigen by preparative polyacrylamide gel electrophoresis.

E. coli BL21 (DE3) carrying the fusion protein construct was grown at 37°C in 2x TY (containing 200ug/ml ampicillin) to $A_{600}=0.8$, at which point IPTG was added to 1mM to induce T7 RNA polymerase (which transcribes message for the fusion protein) and growth continued for a further 3 hours. The cells were pelleted, resuspended in 0.01 volume of 1x loading buffer, sonicated, heated at 90°C for 5 minutes and the proteins (from approximately 50ml original culture volume) resolved on a double thickness SDS polyacrylamide gel (see section 3.23.1). The T7 gene 10 - Oct-1 fusion protein was visualised by briefly staining the gel with 0.05% Coomassie blue in water. The gel slice containing the fusion protein was excised.

3.26.2 Immunisation of rabbit.

Gel slices were mixed with an approximately equal volume of PBS and ground with a pestle and mortar. For the first injection this mixture was emulsified with Freund's complete adjuvant (see Harlow and Lane (1988)). For subsequent injections, slices were emulsified with Freund's

Methods

incomplete adjuvant. Approximately 10ml of blood (for pre-immune serum) was taken from the ear New Zealand White rabbit before the first injection. The rabbit was given 3 injections of 50-100ug of protein at 2 week intervals. Bleeding the rabbit took place 2 weeks after the third injection. Blood was allowed to clot at room temperature for about 6 hours, the clot separated from the side of the tube with a needle and then tubes centrifuged at 2000rpm for 5minutes. The supernatant (serum) was taken and stored in aliquots at -20°C.

3.26.3 Affinity purification of antiserum.

This procedure was adapted from Harlow and Lane (1988). Extract from E. coli expressing the T7 gene 10 - Oct-1 fusion protein was prepared as described in section 3.26.1. Extract containing approximately 100ug of fusion protein was loaded across the entire width of, and run on a SDS polyacrylamide protein gel (see section 3.23.1). The gel was electroblotted onto a nitrocellulose filter (see section 3.24).

A narrow strip was cut from the edge of the filter and the fusion protein detected with the polyclonal antiserum (see section 3.26.2). The section of filter containing the fusion protein was then cut out, and incubated with 3ml of antiserum mixed with 3ml of TBS (20mM Tris.HCl pH7.6, 137mM sodium chloride) on a slow shaker for 4 hours. The filter was washed in 3 changes of TBS over 15 minutes, and then

Methods

bound antibody removed by incubating the filter in 3ml of 100mM glycine (pH2.5) for 5 minutes. The glycine solution (containing affinity purified antibodies) was mixed with 300ul of 1M Tris.HCl (pH8), and aliquots stored at -20°C.

3.27 Sucrose Density Gradient Centrifugation.

Sucrose density gradient centrifugation was used to fractionate oocyte protein extract. Oocyte protein extract was prepared as described for band shift assay protein extract (see section 3.31.1) except that the glycerol in extract buffer was replaced by 8% (w/v) sucrose.

An 11ml 30% (w/v) to 10% sucrose gradient was prepared in an ultracentrifuge tube. The sucrose solutions were made in band shift assay extract buffer, except with no glycerol (50mM Tris.HCl pH8, 50mM KCl, 0.1mM EDTA, 5mM MgCl₂, 2mM DTT). The gradient was cooled to 4°C and 500ul (equivalent to 100 oocytes) of oocyte extract layered on top of the gradient. The gradient was centrifuged in a swing-out rotor at 37000rpm for 45 hours at 4°C.

The gradient was fractionated from the bottom of the tube by inserting a hollow needle attached to a peristaltic pump, and pumping out approximately 20 equal fractions. The fractions were divided before storage, some being stored at -20°C (for use in Western blots and immuno-detection of proteins) and some stored in liquid nitrogen (for use in band shift assays).

Methods

3.28 Preparation of Nuclear Extract from Xtc Cells.

This method (adapted from a method used to prepare nuclei from Xenopus erythrocytes, pers. comm., M. T. Vlad) was used to prepare extract from nuclei of X. laevis Xtc cells, for use in Western blotting and immunodetection of proteins.

Subconfluent Xtc cells (see section 3.1.1) were scraped from the surface of a 75cm² tissue culture flask in 10ml of PBS. The cell suspension was centrifuged at 3000rpm for 5 minutes and suspended in 200ul of suspension buffer (0.25M sucrose, 0.1M KCl, 0.1M Tris.HCl pH7.4, 0.5mM CaCl₂), and then cooled on ice. The cells in this suspension were counted using a haemocytometer slide. Half of this suspension (whole cell extract) was stored at -20°C. The remainder of this suspension was mixed with 10ml of lysis buffer (0.25M sucrose, 0.1M KCl, 0.1M Tris.HCl pH7.4, 0.5mM CaCl₂, 0.05% saponin (Sigma)). The suspension was transferred to room temperature for 3 minutes and then centrifuged at 3000rpm for 5 minutes. The pellet (nuclear extract) was suspended in 100ul of suspension buffer. The nuclear extract was checked microscopically for complete cell lysis and the nuclei counted using a haemocytometer slide. The yield of nuclei was at least 90%.

The nuclear extract was checked as follows. 10ul of extract (whole and nuclear) was mixed with 500ul of NAE (0.3M sodium acetate pH6.5, 1mM EDTA) and the suspension extracted with an equal volume of neutral phenol. Nucleic

Methods

acids were precipitated from the aqueous phase by the addition of 2 volumes of ethanol. After at least 30 minutes at -20°C the nucleic acids were recovered by microfugation for 10 minutes, and resuspended in 10 μl of TE. Nucleic acids were resolved on a 1% non-denaturing agarose gel (see section 3.5.1). Whole extract contained both chromosomal DNA (nuclear) and ribosomal RNA (cytoplasmic). Nuclear extract contained only chromosomal DNA.

Before running on an SDS-protein gel (see section 3.23.1) extract was mixed with an equal volume of 2x loading buffer, and sonicated to shear chromosomal DNA.

3.29 Methods For Radiolabelling DNA And RNA.

3.29.1 Nick Translation.

DNA fragments isolated from low melting point agarose gels were labelled by nick translation for use as probes in Southern and Northern blotting. A 20 μl reaction was assembled in 1x NTB (50mM Tris.HCl (pH7.2), 10mM MgSO_4 , 0.1mM DTT, 50 $\mu\text{g}/\text{ml}$ BSA) containing 100 ng of DNA, 1mM 'cold' dNTPs (dATP, dCTP, dTTP), 40 μCi of α - ^{32}P -dGTP, 1ng of deoxyribonuclease I (Sigma) and 10 units of *E. coli* DNA polymerase I (Amersham). The reaction was allowed to proceed at 14°C for 3 to 4 hours after which 0.2ml of NaE was added. The mixture was run through a 10 ml sephadex G-50 gel filtration column in TE buffer. The higher molecular weight labelled DNA passes through this column faster than unincorporated nucleotides, allowing separation

Methods

of the labelled DNA from unincorporated nucleotides. Labelled DNA (first peak of radioactivity) was collected from the column, heated at 100°C for 5 minutes, cooled on ice, and then added to Southern/Northern blot hybridisation buffer.

3.29.2 End-labelling DNA with γ -³²P-ATP and T4 polynucleotide kinase.

Single stranded oligonucleotides (for use in screening oligonucleotide directed mutants and in primer extension assays) and duplex oligonucleotides (for use in band shift assays) were labelled using this method. Synthetic oligonucleotides are supplied dephosphorylated at the 5' end, and can therefore be labelled by kinasing with γ -³²P-ATP.

A 15 μ l reaction was assembled in 1x C buffer (66mM Tris.HCl (pH7.6), 6.6mM MgCl₂, 10mM DTT) containing 4 to 10 ng of oligonucleotide, 60 to 100 uCi of γ -³²P-ATP and 20 units of T4 polynucleotide kinase. The reaction was incubated at 37°C for 1 hour after which 0.2ml of NaE was added. The mixture was run through a sephadex G-50 gel filtration column in TE buffer, in order to separate labelled oligo from unincorporated nucleotide. Fractions containing labelled DNA were collected and DNA precipitated by the addition of sodium acetate (pH6.5) to a final concentration of 0.3M and 2 volumes of ethanol. DNA was recovered by microfugation for 10 minutes, and resuspended

Methods

in TE buffer.

Essentially the same method was used to end label restriction fragments of DNA from which the 5' phosphate had been removed with calf intestinal alkaline phosphatase (essentially as described in section 3.11.2).

3.29.3 Labelling DNA Fragments by End-Filling.

Up to 1 μ g of restricted DNA (ends of DNA with 5' overhangs) was labelled by filling in the overhang with a nucleotide mix containing one radio-labelled nucleotide. A reaction was assembled in 1x TM buffer (10mM Tris.HCl pH7.5, 5mM MgCl₂) containing DNA, nucleotide mix and 10 units of Klenow fragment of DNA polymerase. The nucleotide mix contained 40uCi of an appropriate α -³²P-dNTP, and the other 3 dNTPs (unlabelled) to give a final concentration of 1mM. The reaction was incubated at room temperature for 30 minutes. Labelled DNA was separated from unincorporated nucleotides by sephadex G50 gel filtration, and recovered by ethanol precipitation.

This method was routinely used to prepare radiolabelled DNA markers from lambda DNA (CI857) digested with Eco RI and Hind III, and from pRR322 digested with Hpa II.

3.29.4 ³²P Labelled Synthetic RNA Probes.

High specific activity RNA probes were prepared (by transcribing subclones in the presence of an α -³²P-NTP) for use in RNase protection assays.

Methods

Subclones in the vector pBluescript SK- (Stratagene) were linearised at an appropriate polylinker restriction site and transcribed in a 20ul reaction containing 1x transcription buffer (40mM Tris.HCl pH8, 6mM MgCl₂, 20mM DTT), 60uCi α -³²P-UTP, 1mM 'cold' NTPs (ATP, GTP, CTP), 62.5ug/ml BSA, 10 units human placental RNase inhibitor (Pharmacia), 100 to 200ng template (linearised DNA), and 10 units of T7 or T3 RNA polymerase (New England Biolabs). The reaction was incubated at 37°C for 30 minutes, and then 5 units of RNase-free DNase I (Boehringer) added, and the reaction incubated at 37°C for a further 10 minutes. The reaction was stopped by the addition of 100ul NAE (0.3M sodium acetate pH6.5, 1mM EDTA), extracted with an equal volume of neutral phenol, and then RNA precipitated from the aqueous phase by the addition of 2 volumes of ethanol. After at least 30 minutes at -20°C RNA was recovered by microfugation for 10 minutes and resuspended in S1 hybridisation buffer (see section 3.11).

Incorporation of label into probe was measured by the following procedure. 0.5% of the probe (in S1 hybridisation buffer) mixed with 5ml of ice cold 10% (w/v) trichloroacetic acid (TCA) and 100ug of salmon sperm DNA. The mixture was incubated on ice for 15 minutes, and then filtered through a glass fibre filter, using a millipore vacuum filtration apparatus. The filter was washed twice in 10ml of 10% (w/v) TCA and twice in 10ml ethanol, dried at room temperature, and then precipitated radioactivity

Methods

measured in a scintillation counter. TCA precipitates the RNA (but not unincorporated nucleotide) which is then retained on the glass fibre filter. 2×10^5 cpm of RNA was used in each RNase protection assay. Typically, 50 times this amount was obtained from each transcription reaction.

3.30 Preparation of Synthetic RNA for Microinjection into Oocytes.

Subclones in the vector pBluescript SK- (Stratagene) were linearised at an appropriate polylinker restriction site and transcribed in a reaction containing 1x transcription buffer (40mM Tris.HCl pH8, 6mM MgCl₂, 20mM DTT), 0.5mM ATP, 0.5mM UTP, 0.5mM CTP, 0.05mM GTP, 62.5ug/ml BSA, 10 units human placental RNase inhibitor (Pharmacia), 200 to 500ug template (linearised DNA), 0.5mM RNA Cap structure analogue (m⁷G(5')ppp(5')G, from New England Biolabs) and 10 units of T7 or T3 RNA polymerase (New England Biolabs). The reaction were incubated at 37°C for 1 hour, then GTP (to a final concentration of 0.5mM) and 10 units of T7 or T3 RNA polymerase added. The reaction was incubated for a further 30 minutes at 37°C. The reaction was stopped by the addition of 200ul of NAE (0.3M sodium acetate pH6.5, 1mM EDTA), extracted with an equal volume of neutral phenol, the aqueous phase extracted with chloroform, and nucleic acids precipitated from the aqueous phase by the addition of 2 volumes of ethanol. After at least 30 minutes at -20°C nucleic acids were recovered by microfugation for 10

Methods

minutes and resuspended in 5 to 10ul of TE buffer. RNA concentration was estimated by electrophoresing a fraction of the product of the transcription reaction on a non-denaturing agarose gel (see section 3.5.1) and viewing the gel under UV light. RNA prepared in this way was microinjected into the cytoplasm of oocytes, and translation products labelled with ^{35}S -methionine (see section 3.3).

3.31 Preparation of Protein Extract for Band Shift Assays.

The following procedure was used to prepare protein extracts for use in band shift assays (see section 3.32), and extracts prepared in the same way were also used for SDS-protein acrylamide gel electrophoresis. All extracts were made in extract buffer: 50mM Tris-HCl pH8, 50mM KCl, 0.1mM EDTA, 5mM MgCl_2 , 10ug/ml aprotinin (Boehringer), 10ug/ml leupeptin (Boehringer), 25% glycerol.

3.31.1 From *Xenopus laevis* oocytes, eggs and embryos.

Oocytes and embryos were homogenised on ice in 5ul of extract buffer per oocyte/embryo, microfuged at 4°C for 10 minutes and the supernatant snap frozen in liquid nitrogen. Enucleated oocyte extract was prepared in the same way from oocytes from which the nuclei had been removed by manual dissection (the animal pole of the oocyte was pierced with the tip of a fine needle, and the oocyte gently squeezed with a pair of watchmakers forceps to expel the nucleus

Methods

with little loss of cytoplasm). Isolated nuclei were picked up with a Gilson pipette in 5ul of 1x barth-X (the medium in which oocytes were incubated), glycerol added to 25% (v/v), and the extract snap frozen in liquid nitrogen. This extract was tested in band shift assays, but was not active (either because of leakage of nuclear contents before being picked up, or because the barth-X does not maintain binding proteins in an active form). However, nuclear extract was used for SDS-protein gel electrophoresis.

3.31.2 From adult X. laevis tissues.

Minced tissues were homogenised on ice in 2-3 volumes of extract buffer, microfuged at 4°C for 10 minutes and the supernatant snap frozen in liquid nitrogen.

3.31.3 From cultured cells.

Subconfluent cells were washed in phosphate buffered saline (PBS) and scraped from the surface of the 75cm² flask in 10mls of ice cold PBS, pelleted and resuspended in 30ul of extract buffer per flask. The suspended cells were homogenised on ice, microfuged at 4°C for 10 minutes and the supernatant snap frozen in liquid nitrogen. When levels of binding proteins were to be compared the protein concentration of the extract was measured (Biorad protein detection kit) and equalised at 5mg/ml.

Methods

3.31.4 From E. coli expressing fusion proteins.

E. coli BL21 (DE3) and E. coli TG2 carrying fusion protein constructs were grown in 2x TY (containing 200ug/ml ampicillin) to $A_{600} = 1.5$, pelleted and resuspended in 0.01 volumes of extract buffer. The suspended cells were sonicated on ice, microfuged at 4°C for 10 minutes and the supernatant snap frozen in liquid nitrogen.

3.32 Band shift assays.

3.32.1 Preparation of probes.

Complementary synthetic oligonucleotides (25mers) were mixed at a final concentration of 1mg/ml in 1x medium salt restriction digest buffer (10mM Tris.HCl pH7.4, 10mM MgSO₄, 50mM NaCl, 1mM DTT), heated at 85°C for 10 minutes, and then left to cool to room temperature in an eppendorf tube floating in a 1 litre beaker of water at 85°C. Annealed oligonucleotides were then diluted to a concentration of 2ug/ml in TE buffer. 10ng of duplex oligonucleotide was end labelled with 100uCi of γ -³²P-ATP and T4 polynucleotide kinase (oligonucleotides are supplied dephosphorylated) as described in section 3.29.2.

3.32.2 Assays.

25ul binding reactions were carried out on ice for 20 minutes in binding buffer (45mM KCl, 15mM HEPES pH7.9, 5mM spermidine, 1mM MgCl₂, 1mM dithiothreitol, 0.5mM PMSF, 0.1mM EDTA, 7% glycerol) containing 2ug of pAT153, 1ug of

Methods

salmon sperm DNA, 20ng of non-specific duplex oligonucleotide (ACAGACCGAAGCTTAGCT), 0.3ng of duplex oligonucleotide probe, up to 5ul of protein extract (see section 3.21) and (when added) 1ul of antiserum. For competition analysis the non-specific duplex oligonucleotide probe was replaced by 20ng of specific duplex oligonucleotide.

The reactions were electrophoresed on a 5% polyacrylamide gel (29:1 bis) in Tris-borate buffer (22.5mM Tris-borate pH8, 0.5mM EDTA) at 200 volts and 4°C for 2 hours. The gel was fixed, dried and exposed to X-ray film with an intensifying screen for 2-12 hours at -70°C.

Results and Discussion

RESULTS AND DISCUSSION.

Chapter 4.

Homologues of human Oct-1 occur in the *X. laevis* genome, and in *X. laevis* oocyte RNA.

Introduction.

The clone pBS-Oct-1⁺ (see Sturm et al (1988A), gift of Winship Herr) contains the entire human Oct-1 cDNA in the vector pBSM13⁺ (Stratagene). An Eco RI digest releases the Oct-1 cDNA from this clone as two fragments of approximately equal size (about 1.2 and 1.3 kb). These two fragments were isolated together from an agarose gel, and the mixture labelled for use as a probe for a *X. laevis* genomic DNA Southern blot, and for a *X. laevis* A⁺ selected RNA Northern blot. These were carried out as a preliminary to screening for *X. laevis* cDNA homologues of human Oct-1 to ensure that homologues existed.

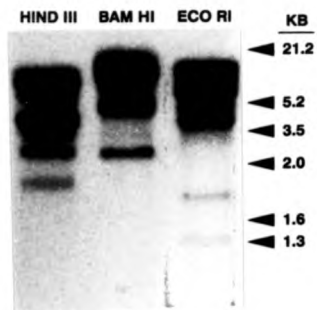
4.1 *X. laevis* genomic Southern blot.

X. laevis genomic DNA was digested with Eco RI, Bam HI and Hind III, electrophoresed on an agarose gel, Southern blotted to a nitrocellulose filter and then hybridised to labelled human Oct-1 cDNA. The resulting autoradiograph is shown in figure 7. Homologues of human Oct-1 are clearly seen at relatively high stringency (final washing condition were 0.2x SSC at 42°C). The many hybridising bands may represent several genes. There are several mammalian POU domain-

Figure 7.

Southern blot of X. laevis genomic DNA probed with human Oct-1 cDNA.

10ug of X. laevis genomic DNA prepared from the blood of a single individual was digested with the restriction endonucleases indicated, electrophoresed on a 0.8% agarose gel, and the gel Southern blotted to a nitrocellulose filter. The filter was hybridised to 50ng of human Oct-1 cDNA labelled by nick translation. Hybridisation was in 6xSSC and 50% formamide at 42°C, and final washing conditions were 0.2x SSC at 42°C. An autoradiograph after an eight hour exposure to the filter is shown. Lambda DNA digested with Eco RI and Hind III was run as a marker, and the position and size of marker bands (determined by ethidium bromide staining of the gel before transfer) is shown.



Results and Discussion

containing genes, and consequently the Oct-1 POU domain present in the probe may hybridise to the POU domains of a family of Xenopus laevis genes. Alternatively the many hybridising bands could indicate that the Oct-1 gene is large with several introns. Also, X. laevis is tetraploid with respect to a theoretical ancestor, and consequently there are probably two copies of a X. laevis Oct-1 gene. The signal strength varies (particularly in the Hind III track), with some bands being weaker than those above or below. This may indicate that sequences with varying degrees of homology to human Oct-1 are present.

4.2 X. laevis A+ selected RNA Northern blot.

I wanted to isolate a X. laevis cDNA homologue of human Oct-1. Two X. laevis cDNA libraries were available in the laboratory: adult liver (made by A. R. Brooks) and mature folliculated oocyte (gift of J. Shuttleworth). Human Oct-1 is ubiquitous, and consequently both libraries should be suitable for the isolation of X. laevis Oct-1. However, a Northern blot of X. laevis A+ selected RNA from these two tissues was probed with human Oct-1 cDNA to check that homologous transcripts were expressed in these tissues in order to determine which of the two libraries to screen.

X. laevis A+ selected RNA from oocytes and liver was electrophoresed on a 1.5% formaldehyde-agarose gel, Northern blotted to a nitrocellulose filter, and the filter hybridised to labelled human Oct-1 cDNA. The resulting autoradiograph is

Figure 8.

Northern blot of X. laevis RNA probed with human Oct-1 cDNA.

10ug of X. laevis A+ selected RNA (from the tissue indicated) was electrophoresed on a 1.5% formaldehyde-agarose gel, the gel Northern blotted to a nitrocellulose filter, and the filter hybridised to 50ng of human Oct-1 cDNA labelled by nick translation. Hybridisation and washing conditions were the same as those described for the Southern blot in figure 7. An autoradiograph after a five day exposure to the filter is shown. The position of ribosomal RNA bands is indicated, and was determined by ethidium bromide staining of the gel before transfer.

APPROX. ★
10KB

OOCYTE A+
LIVER A+

▲ 28 S (4.5KB)

▲ 18 S (1.9KB)



Results and Discussion

shown in figure 8. Hybridisation to low molecular weight RNA is seen in the oocyte RNA track (probably non-specific) and also hybridisation to a RNA species of an estimated size (from the position of ribosomal RNA bands) of around 10kb. This indicates that a large transcript from an Oct-1 homologous gene is expressed in X. laevis oocytes. The transcript is of low abundance (a five day exposure of the filter is shown). The size of the human Oct-1 cDNA is around 2.5kb, and so the homologous X. laevis transcript is unexpectedly large. However, large human Oct-2 transcripts have been reported (approximately 6kb, see Muller *et al* (1988)) and human Oct-1 message is a similar size to the transcript detected in X. laevis oocyte RNA (W. Herr, pers. comm.). Much of the excess size over the published cDNA sequence is accounted for by the fact that the published cDNA sequence does not have the genuine 3' end, and the genuine human transcript has a very long 3' untranslated region. No hybridisation of human Oct-1 to X. laevis liver RNA is detected. This result does not necessarily mean that the X. laevis homologue of Oct-1 is not expressed in liver (which would be unlike human Oct-1), but may reflect a difference in abundance between oocytes and liver. As will be shown later, Oct-1 protein is stored in large amount in X. laevis oocytes, and the message for Oct-1 could have a higher abundance in oocytes. As a consequence of this result the X. laevis mature folliculated oocyte cDNA library was selected to screen for homologues of human Oct-1.

Results and Discussion

Chapter 5.

Isolation of *Xenopus laevis* Oct-1 cDNA clones from an oocyte library.

5.1 Library screening.

A cDNA library (gift of Dr. J. Shuttleworth) prepared with A+ selected RNA from *X. laevis* oocytes, which had been matured in vitro with progesterone, was screened using human Oct-1 as a probe. Eco RI linkers had been added to the cDNA, and the linked cDNAs ligated into the Eco RI site of the vector lambda Zap (Stratagene). The library contained 4×10^5 independent clones, with inserts in the size range 1 to 2.5 kb, and I received an aliquot of amplified phage stock with a titre of 1×10^{10} pfu/ml. The layer of follicle cells which surrounds the oocytes had not been removed, and consequently cDNAs from these cells are represented in the library.

A total of 1.1×10^6 phage from the oocyte cDNA library were screened using the human Oct-1 cDNA probe described in chapter 4 (for use in Southern and Northern blots). Screening was as described in the Methods section, taking replica filters from each plate. Twelve replica positive plugs were taken from the first screen, and rescreened until single positive plaques were obtained. Three screens were required to obtain single positive plaques. An example of a third round screen is shown in figure 9. Eleven of the twelve positives were shown to be genuine positives on rescreening. These eleven plaques were then purified by plating-out, and picking single plaques without rescreening.

Figure 9.

A third round screen of a positive cDNA clone from the X. laevis oocyte cDNA library.

Phage from a plug containing a plaque which hybridised with human Oct-1 cDNA were plated out at low density onto E. coli BB4, and replica plaque lifts taken from the plate. The replica filters were hybridised to human Oct-1 cDNA labelled by nick translation. Hybridisation was at 42°C in 6x SSC and 50% formamide, and final washing conditions were 42°C in 1x SSC. An autoradiograph of the replica filters (in the same orientation with respect to the plate from which the lifts were taken) is shown. Plugs containing a single replica positive were taken following this third round screen.

FIRST
LIFT



SECOND
LIFT



Results and Discussion

5.2 Automatic excision of the prospective Oct-1 cDNA clones from the phage vector and preliminary restriction analysis.

The eleven prospective Oct-1 cDNA clones were converted to the plasmid form (in pBluescript sk-) by automatic excision from the phage vector, according to the manufacturer's instructions. Two single colonies were picked and used to prepare miniprep DNA, which was then digested with Eco RI to release the insert. Two colonies were analysed for each clone because there is some evidence that gross rearrangement can occur during excision (J. Shuttleworth, pers. comm.). However, the duplicate colonies gave identical products on Eco RI digestion, indicating that no rearrangement had occurred. Eco RI digestion released single fragments. Seven clones contained an approximately 2.5kb insert, two clones an approximately 2.3kb insert, and two clones an approximately 1.9kb insert. This indicated that three individual clones had been isolated, and this was confirmed since representatives of the same clone gave the same Pst I digestion pattern (there are multiple Pst I sites, the map is described later). Consequently representatives of the three individual clones (clones 3, 6 and 16) were selected for further analysis.

5.3 Detailed restriction analysis of clones 3, 6 and 16.

A more detailed restriction analysis of the three individual positive clones was carried out in order to be able to determine a sequencing strategy, and to indicate if the three

Results and Discussion

clones were related.

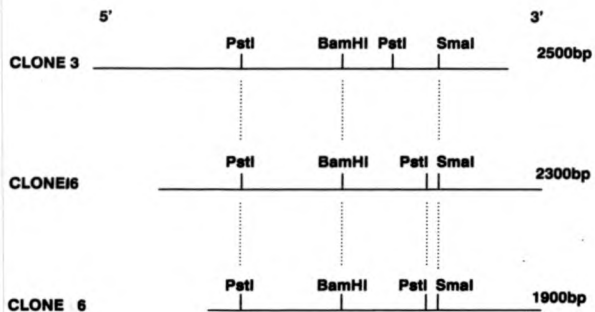
Standard single and double restriction enzyme digests were used to determine the preliminary restriction enzyme maps shown in figure 10. It was determined that none of the three clones had sites for Hind III, Sal I, Kpn I, Xba I, Bgl II, and Sac I. Each clone had a single site for Bam HI and Sma I, and the position of these sites was determined by double digests with a non-cutter located at one end of the clone, in the polylinker. Pst I digestion of clones 6 and 16 released two fragments, indicating the presence of two sites, whose position was mapped by double digests (with Bam HI, Sma I and non-cutters in the polylinker at either end of the cDNA). Pst I digestion of clone 3 released six fragments, indicating six sites for Pst I. One fragment was approximately 0.9kb, and the other 5 fragments small, in the size range 100 to 250 bases. Consequently, the sites could not be mapped completely by double digestion. However, Bam HI cuts the 0.9kb Pst I fragment, and consequently the fragment was determined to be centred around the Bam HI site, and flanked by two internal Pst I sites. The position of the two Pst I sites was determined by a Sma I - Pst I double digest. This lead to the Bam HI, Sma I and partial Pst I maps shown in figure 10. The three clones are shown aligned by apparently common restriction enzyme sites. It should be noted that the 5' most Pst I site shown in each clone was shown by sequencing to be two very close Pst I sites, and that clone 3, in fact, contains nine Pst I sites. This discrepancy (restriction

Figure 10.

Partial restriction maps of clones 3, 6 and 16 aligned to indicate conservation of some sites.

NB. Restriction digest indicated four additional Pst I sites in clone 3 which were not mapped. Later sequence analysis revealed that clone 3 in fact contains nine Pst I sites, and clone 16 three Pst I sites. The complete maps are shown later.

5' and 3' ends of the clones were tentatively assigned on the basis of the lack of hybridisation of clone 6 to the 5' end of human Oct-1 (see figure 11).



Results and Discussion

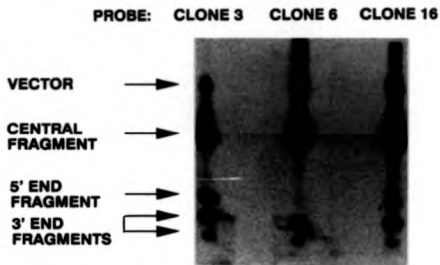
analysis indicated six sites) arose because of the aforementioned 'doublet' of Pst I sites, and because there is a 'triplet' of very close Pst I sites. The conservation of all Bam HI, Pst I and Sma I sites between clones 6 and 16 indicated that these clones were derived from the same message, clone 6 simply being shorter at one end, probably the 5' end as a result of premature termination of cDNA synthesis (cDNA synthesis was primed from the 3' polyA⁺ tail with oligo dT). The conservation of Bam HI and Sma I sites and one Pst I site between clone 3 and the other two clones indicates that this clone is related to, but distinct from the other two clones.

The suggestion that clones 6 and 16 were equivalent, except that clone 16 was longer at the 5' end was tested further. A Bam HI - Pst I double digest of pBS-Oct-1+ (human Oct-1 cDNA in pBSM13+, see chapter 4) generates five Oct-1 fragments: a central approximately 1.3kb fragment; a 5' approximately 450bp fragment; and three 3' fragments (one of approximately 300bp, and two of approximately 150bp). A map is shown in figure 11A. Such digests were Southern blotted and hybridised to labelled clones 3, 6 and 16. The result is shown in figure 11B. Clones 3 and 16 detect all fragments. Clone 6 does not detect the 5' end fragment. Consequently it was concluded that clones 6 and 16 were equivalent, clone 6 simply representing premature termination of cDNA synthesis from the same transcript. On the basis of the lack of hybridisation of clone 6 to the 5' end of human Oct-1, and alignment of the

Figure 11.

Clone 6 does not hybridise to a fragment from the 5' end of human Oct-1.

- (A) A Bam HI - Pst I map of the human Oct-1 cDNA from pBS-Oct-1+ (Sturm et al (1988A)).
- (B) Three Bam HI - Pst I digests of pBS-Oct-1+ were electrophoresed on a 1% agarose gel and the gel Southern blotted to a nitrocellulose filter. The filter was cut into strips, and one strip hybridised to the Eco RI insert from clone 3, 6 or 16, labelled by nick translation. The hybridisation and washing conditions were as described for the Southern blot in chapter 4. The positions of fragments from human Oct-1 cDNA and the position of the vector fragment is indicated.

A**B**

Results and Discussion

three clones by apparently conserved sites, the 5' ends of the three clones were tentatively assigned.

Clones 3 and 16 were, therefore, selected for sequence analysis. It was intended to sequence from M13 subclones, and to facilitate this restriction maps were extended to include Stu I and Pvu II sites. These sites were mapped in the same way as Bam HI, Sma I and Pst I sites, using single and double digests. The position of these sites is shown in later maps describing the sequencing strategy.

Results and Discussion

Chapter 6.

Sequence analysis of Oct-1 homologues.

6.1 Sequencing clones 3 and 16.

Clones 3 and 16 were completely sequenced. The sequencing strategy is shown in figures 12 and 13. Sequencing (by the dideoxy chain termination method) was largely from single stranded template, derived from M13 subclones, with the M13 17mer sequencing primer. The position of Sma I, Bam HI, Pvu II, and Stu I sites used for subclones was determined by restriction mapping. The position of Pst I sites used for subclones was determined partly by restriction mapping, and partly from the sequence of other subclones. Three 17mer oligonucleotides were made using the sequence of subclones and used as sequencing primers. In one case the primer was used to sequence from a double stranded plasmid template (since the sequence obtained from single stranded M13 templates indicated that the primer also hybridised to a site in the M13 sequence). During sequencing all restriction sites were overlapped to ensure that sites determined by restriction mapping were not in fact 'doublets' of two sites close together. This turned out to be the case for some Pst I sites, where sites were too close for the small intervening fragment released on digestion to be seen on a agarose gel. Any ambiguities in the sequence were resolved by sequencing both strands.

The complete nucleotide sequence of clone 3 is shown in figure 14. Translation of the longest open reading frame,

Figure 12.

Sequencing strategy for clone 3.

E = Eco RI
F = Fst I
V = Fvu II
T = Sst I
B = Bam HI
S = Sma I

Sequencing was largely from the subclones indicated by the boxes, made in M13mp18 or M13mp19. The arrows indicate the direction and extent of sequencing. Sequencing from these subclones was from the M13 17mer sequencing primer on single stranded template. Two oligonucleotides were made to the sequence of clone 3 (17mers, position indicated) and sequencing from these primers was carried out as follows: from oligo 2, using the complete Eco RI fragment in M13; from oligo 1, using the complete Eco RI fragment in pBluecript SK- (ie. sequencing on a double stranded plasmid template).

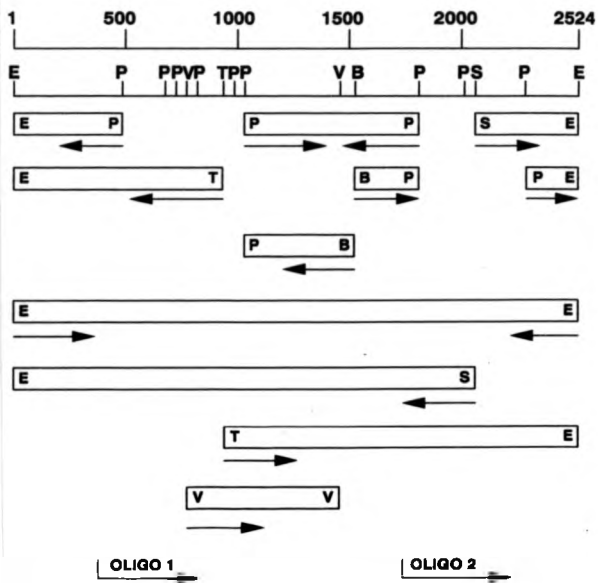
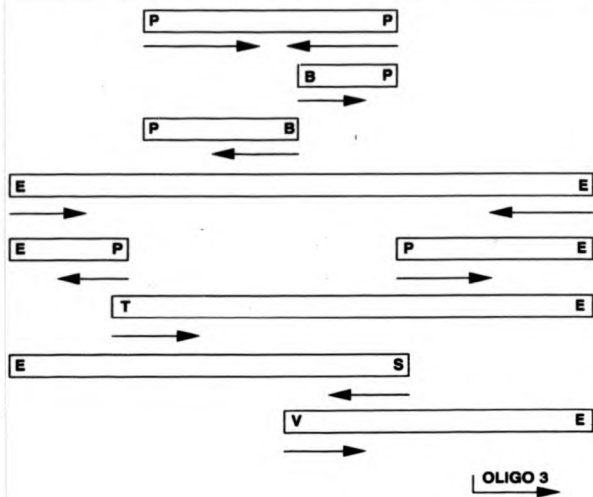
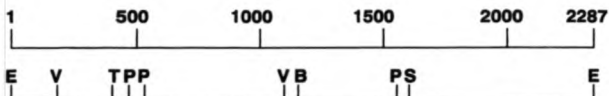


Figure 13.

Sequencing strategy for clone 16.

E = Eco RI
P = Pst I
V = Pvu II
T = Stu I
B = Bam HI
S = Sma I

Sequencing was largely from the subclones indicated by the boxes, made in M13mp18 or M13mp19. The arrows indicate the direction and extent of sequencing. Sequencing from these subclones was from the M13 17mer sequencing primer on single stranded template. One oligonucleotide was made to the sequence of clone 16 (17mer, position indicated) and sequencing from this primer was carried out on the Pst I - Eco RI fragment in M13.



Results and Discussion

from the first in frame methionine is shown. A stop codon occurs three codons before this methionine, indicating that this is the authentic translation start point. The clone has no poly A+ tail, which could be a result of premature second strand cDNA synthesis. The clone probably represents a full-length coding sequence.

The complete nucleotide sequence of clone 16 is shown in figure 15. Translation of the longest open reading frame is shown. There are no in frame stop codons upstream of the first methionine, and as will be confirmed in a later comparison with human Oct-1, this clone probably represents an incomplete reading frame. This clone has a polyA+ tail, and two potential polyA+ addition signals are shown. Consequently this may represent the authentic 3' end of the message.

An alignment of the nucleotide sequences is shown in figure 16. The two cDNAs show 93% identity, within the region of overlap. As predicted from the restriction analysis (by conservation of some sites, and hybridisation to a human Oct-1 digest) clone 3 is longer at the 5' end, and clone 16 longer at the 3' end. As will be described in the following section, clones 3 and 16 represent genuine homologues of human Oct-1, and henceforth will be described as follows: clone 3 = Xl-Oct-1A (for Xenopus laevis Oct-1, A) and clone 16 = Xl-Oct-1B.

Figure 14.

Complete nucleotide sequence of clone 3 (Xl-Oct-1A).

Translation from the first in frame methionine is shown. The position of an in frame stop codon two amino acids before the first in frame methionine is shown.

[illegible]

Figure 15.

Complete nucleotide sequence of clone 16 (X1-Oct-1B).

Translation of the longest open reading frame is shown. The first in frame methionine is boxed. There are no in frame stop codons 5' of the first in frame methionine. The location of two potential polyA addition signals (AATAAA) is indicated.

[illegible]

Figure 16.

Nucleotide sequence alignment of clone 3 (Xl-Oct-1A) and clone 16 (Xl-Oct-1B).

Clone 3 is shown above clone 16. Vertical lines represent a conserved nucleotide, and gaps in the sequence are made for maximum alignment. Clone 3 is longer at the 5' end, and clone 16 longer at the 3' end. Within the region of overlap there is 93% homology.

[illegible][illegible]

Results and Discussion

6.2 Comparison of *X. laevis* Oct-1 protein sequences with human and chicken Oct-1.

Figure 17 shows a predicted protein sequence alignment for human Oct-1, chicken Oct-1, Xl-Oct-1A and Xl-Oct-1B.

The C terminal ends of Xl-Oct-1A and Xl-Oct-1B match the C terminal ends of chicken and human Oct-1. The N terminal end of Xl-Oct-1A extends further than chicken Oct-1, and as mentioned earlier contains a stop codon before the first in frame methionine. Consequently Xl-Oct-1A probably contains a complete open reading frame. Xl-Oct-1B is 83 amino acids shorter than Xl-Oct-1A at the N terminal end, and the reading frame contains no stop codon before the first in frame methionine. Consequently this clone probably does not represent a complete open reading frame. Within the region of overlap, Xl-Oct-1A and Xl-Oct-1B are 93% homologous in predicted amino acid sequence (and 93% homologous in nucleotide sequence). *Xenopus laevis* is tetraploid with respect to a theoretical ancestor. This high level of homology probably means that the two clones are copies of the same gene, arising from the theoretical genome duplication event.

Conservation between species is strikingly high. Human and chicken Oct-1 are 96% homologous in amino acid sequence throughout the entire length, and identical within the POU specific and homeo boxes. Over a stretch extending from the first in frame methionine of human Oct-1 to the C terminal end, human Oct-1 and Xl-Oct-1A are 85% homologous in amino acid sequence. In the region of overlap human Oct-1 and Xl-Oct-1B

Figure 17.

Predicted amino acid sequence alignment of human Oct-1, chicken Oct-1, X1-Oct-1A and X1-Oct-1B.

The sequences are aligned to human Oct-1.

- = residue conserved with human Oct-1.

Differences to the human sequence are shown.

* = gap in sequence made for maximum alignment.

Human and chicken Oct-1 show 96% homology.

Human Oct-1 and X1-Oct-1A/B show 85% homology.

Human Oct-1 sequence is from Sturm et al (1988A).

Chicken Oct-1 sequence is from Petryniak et al (1990).

HUM 001 SQDESS**AAAAAADSRHMPSTSEFSHESGCHTCTGTGCLDPQKQVFPVCGAISTAGAAQAFGLHL
 CHX 001 *****
 XIA 001 HKLH--SLIQMH-WLS-A-----P-----A--*I--TS--L--
 XIB 000 *****

HUM 068 HQVQLACTSLQAAAQSLHWQSKSHKESGDSQQSPSQSPVQAAIPQTQLMLAGCQITCLTLTPAQQOL
 CHX 052 -----
 XIA 069 -----N-----T-FK--P-EPM-VV-----L-----V--P-A-----H
 XIB 001 *****T-FK--P-EPM-VG-----N-----V--P-A-----H

HUM 138 LLQQAQAQALLAAAVQSHASQSHASAGATISASAATPMTQIPLSQPIQIAQDLQQLQQLQQLHJLQQ
 CHX 132 **-----
 XIA 139 -----E-----
 XIB 095 **-----P-----

HUM 208 FVLVFTYBLQPAQFIISQTPQCGQLLQQAQLQTLFQSQAWLLQSQPSITLTSQPATPTTYAATPI
 CHX 186 -----L-----A-----
 XIA 206 -----L-----
 XIB 122 -----L-----V

POD-SPECIFIC BOX

HUM 278 QTLFQSQSTPKRIDTPSLHPSDLKLEQFANTPQSHIKLZPTQCDVGLAKKLYCWHPSQTTISHPKA
 CHX 256 P-----
 XIA 276 -----
 XIB 191 -----

HUM 348 LMLVFKHRYKPLKAGWDA**KMLGGUSLSSPSALNSP**GICLSHKKERTSYETNIVALKK
 CHX 326 -----GQ-----VS-----
 XIA 346 -----VL-IT--T-TQ-V--GH-N--
 XIB 261 -----IT*-T-TQ-V--GH-N--

POU HOMBO BOX

HUM 414 FLENKQPTSERITHIADQLNNEKEVIRVFCNERQKKRINFPFSGGTSSSPIKAIFFSPPTSLVATPPL
 CHX 394 -----S-----SL-S--NP--S-----
 XIA 416 -----S-----SL-S--**
 XIB 328 -----S-----SL-S--**

HUM 484 VTSAAATLTVSPVLPLTS*AAVTLVSTGYSSTYSMTATVISTAPPASSAVTSPSLSPSPSASASTSR
 CHX 464 -----N-----P-S--TE-----I--
 XIA 486 -----P-A--N-----A-I-SPHIP--TC-S-A-----V-VL-----S--T-AS--
 XIB 396 -----T-A--N-----A-I-GP--P--TC-P-A-----I--VL-----S--T-A--

HUM 553 ASSASSTSTTQTSTPLSSPLGYSQVWYASGLQTAATAAALQGAALPANASLAAMAAGLWPSLMAPS
 CHX 533 -----G-S-----
 XIA 556 -----T-G-----MT-S-M-C-----IB--Y-----G-----
 XIB 466 -----T--N-----MT--S-C-----IB**--Y-----G-----

HUM 623 QFAAGCALLSLNPGYLSGALSALMSWSTLAIYIQAASGSLPITSLDATCHLVFVWAGGAFNIVTAPLP
 CHX 603 -----S-----S-----T-----
 XIA 626 -----P--A-CS-----S-----A-----T-----
 XIB 534 -----T--P--A-CS-----S-----S--P--T-----

HUM 693 LNPQSLILTSNPFVLSAAAAAGHSAFVASLEATSTAESIQHSLPTVASAGAASTTTTAAKAG 759
 CHX 673 -----A-C-----L-----T-----
 XIA 696 -----P-----I--**--ATC-IT--TS-ID--A-N-----S--760
 XIB 604 -----*****--ATC-IT--TL-ID--T-N-----S--665

Results and Discussion

are 85 % similar in amino acid sequence. This level of homology must imply that the proteins have closely related functions. Within the POU specific box human Oct-1 and Xl-Oct-1A are 97% homologous (a two amino acid insertion in Xl-Oct-1A, just before the non-conserved linker) and human Oct-1 and Xl-Oct-1B are identical. Within the POU homeo box Xl-Oct-1A and human Oct-1 are identical, and Xl-Oct-1B and human Oct-1 are 97% homologous (two amino acid substitutions). Consequently conservation in the POU homeo and specific boxes is higher than elsewhere in the protein. The differences in the POU homeo and specific boxes (two amino acid differences in each) between Xl-Oct-1A and Xl-Oct-1B are probably not functionally significant, although three amino acid substitutions in the human Oct-1 homeo box do prevent interaction with the viral transactivator, VP16 (see section 1.3.5). Within the non-conserved linker (between POU homeo and specific boxes) human and X. laevis Oct-1 are about 50% similar. This is in line with evidence that the linker has no specific functional purpose (see section 1.3.2).

There is no stop codon in front of the published first in frame methionine of human Oct-1, and this reflects the fact that this is not the authentic translation start site (W. Herr and G. Das, pers. comm.). The Oct-1 reading frame extends further at the N terminal end than published. The current unpublished state of the human Oct-1 N terminal end (which probably extends further than the longest cDNA so far cloned) is illustrated in figure 18. Alternate splicing

Figure 18.

Alternate splicing occurs at the N terminal end of human Oct-1.

The sequence of the N terminal end of human Oct-1, which extends further than the published sequence is shown (W. Herr and G. Das, pers. comm.). An additional exon is inserted in some human Oct-1 cDNA clones by alternate splicing. This splice brings a stop codon into frame, which may be overlooked by ribosomal frameshifting. The sequence of the N terminal end of X1-Oct-1A is aligned to human Oct-1 to show that there is a complete breakdown in homology after the splice site.

X1-Oct-1A MKLHS
 hu Oct-1 QSQSEGGFIDRAILVKIFLMADGGAASQ

X1-Oct-1A SSLIQNH-WLS- A-----*-----P-----
 hu Oct-1 DESSAAAAAAAD SEMNNPSETSKFPMESGD.....

WKSCKSPFPAFLIKLLSVFMESVQKKAFFLYL

ADDITIONAL KION IN
 ALTERNATE SPLICED
 FORM

↓
 +2 bases, results
 in frame-shift
 (stop codon after
 two residues).

- = conserved residue
 * = gap for maximum alignment

Results and Discussion

occurs at the N terminal end, resulting in an insertion of approximately 100 bases in some human Oct-1 cDNAs. The N terminal end of XI-Oct-1A displays no homology to either spliced form of human Oct-1 upstream of this splice site. Consequently, XI-Oct-1A may represent a third alternately spliced form. It should be noted that the presence of the additional exon in human Oct-1 causes the reading frame to include a stop codon two amino acids after the splice site. W. Herr and G. Das (pers. comm.) have evidence to suggest that this stop codon is overlooked by ribosomal frameshifting (for recent example see Belcourt and Farabaugh (1990)).

Results and Discussion

Chapter 7.

Southern blot of *X. laevis* genomic DNA probed with POU domain, Oct-1 specific and Oct-2 specific probes.

X. laevis genomic DNA was digested with restriction enzymes, fractionated through an agarose gel and then Southern blotted to a nitrocellulose filter. DNA was hybridised to labelled probe containing the POU domain of Xl-Oct-1B, or a fragment from the C terminus of Xl-Oct-1B (which would not be expected to hybridise to other POU domain containing genes, and hence termed the Oct-1 specific probe), or a fragment from the N terminus of human Oct-2A (termed the Oct-2 specific probe). The human Oct-2A clone was a gift of P. Matthias (see Muller et al (1988)). The origin of the probes is illustrated in figure 19A, and an autoradiograph of the Southern blot in figure 19B.

Oct-1 specific and POU domain probes give essentially the same hybridisation pattern to Eco RI and Hind III genomic DNA digests. The hybridisation pattern to a Pst I digest is distinct for the two probes, but this is to be expected since Pst I cuts between the two probes in the cDNA, and will therefore do so in the gene. However, it is striking that the Eco RI/Hind III hybridisation pattern is similar, and that the POU domain probe does not react to a large number of fragments, since in mammals there is a large family of POU domain-containing genes. This may indicate that in *X. laevis* there is no large family of POU domain genes, or that if such a family exists the POU domains are sufficiently diverged

Figure 19.

Southern blot of X. laevis genomic DNA probed with POU domain, Oct-1 specific and Oct-2 specific probes.

- (A) Restriction maps of Xl-Oct-1B and human Oct-2A (Muller et al (1988)) showing the fragments isolated for use as Southern blot probes.
- (B) 10 ug of X. laevis genomic DNA digested with the restriction endonucleases indicated was electrophoresed on a 0.8% agarose gel, Southern blotted to a nitrocellulose filter, and then DNA hybridised to the probes indicated, labelled by nick translation. Hybridisation was in 6x SSC and 50% formamide at 42°C, and final washing conditions were 2x SSC at 50°C. Lambda DNA digested with Eco RI and Hind III was run as a marker, and the position of marker bands is indicated.

Results and Discussion

from the Oct-1 POU domain to be not detected even at low stringency (final washing conditions were 2x SSC at 50°C).

The hybridisation of the human Oct-2 specific probe to X. laevis genomic DNA digests gives a 'smudged' appearance. However, at least two high molecular weight fragments are detected in the Xco RI track. This may indicate that the X. laevis genome contains homologues of Oct-2. However, as described above the X. laevis Oct-1 POU domain does not appear to hybridise to other POU domains, which could mean that (unlike mammals) X. laevis Oct-1 and Oct-2 POU domains are significantly diverged. Alternatively, there may not be an Oct-2 homologue. The probe contains a glutamine rich region of Oct-2, and glutamine rich probes hybridise to multiple X. laevis cDNAs at low stringency (J. C. Richardson pers. comm.). This could account for the 'smudged' hybridisation pattern.

Results and Discussion

Chapter 8.

Detection of Xl-Oct-1A and Xl-Oct-1B transcripts in X. laevis oocyte RNA by RNase protection assays.

Introduction.

It was decided to assay the levels of Oct-1 transcripts in oocytes using RNase protection assays, in order to demonstrate that the Oct-1 clones were from X. laevis oocytes and not contaminants from some other source, and to make a rough estimate of the amount present. This method was chosen rather than Northern blot hybridisation because it is more sensitive, and Northern blots required polyA⁺ selected RNA and long exposures to detect Oct-1 transcripts in oocytes, and could not detect Oct-1 transcripts in liver (see chapter 5). Also, the Northern analysis detected a single transcript, whereas RNase protection assays should be able to detect and distinguish Xl-Oct-1A and Xl-Oct-1B transcripts. RNase protection assays do, however, have the disadvantage that if transcripts were to vary, for example by alternate splicing, they would probably not be detected. Whilst this work was in progress it was reported that the levels of X. laevis Oct-1 transcripts, assayed by RNase protection, were roughly constant through early development (M. Perry, pers. comm.).

Results and Discussion

8.1 Subcloning of a fragment of Xl-Oct-1A into a transcription vector, for use in the preparation of synthetic RNA probes.

The 265 bp 5' Bam HI to 3' Pst I fragment from Xl-Oct-1A (see restriction map, figure 12) was isolated and subcloned into pBluescript sk- (Stratagene). Labelled antisense transcripts were made using T7 RNA polymerase from template DNA which had been linearised at the polylinker Sac I site. Including polylinker sequences the probe produced in this way is approximately 360 bases in length. The fragment protected by Xl-Oct-1A should be 265 bases, and the difference in length between probe and protected fragment allows protected probe to be distinguished from undigested probe.

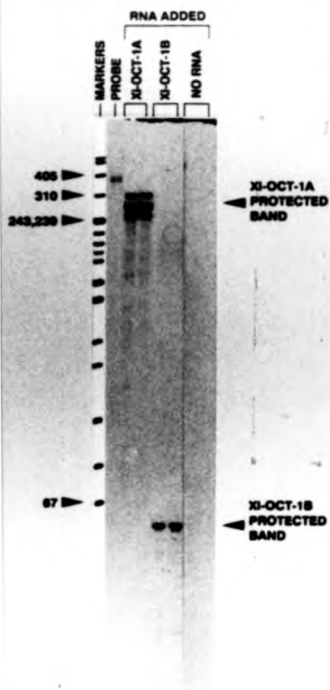
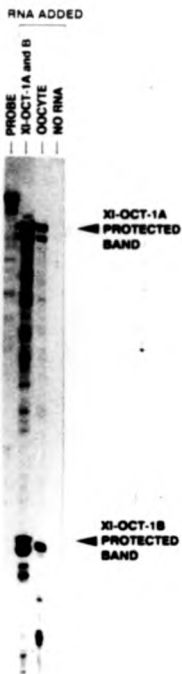
8.2 The probe detects both Xl-Oct-1A and Xl-Oct-1B synthetic RNA.

As a test to check the assay system, assays were set-up with full length unlabelled sense Xl-Oct-1A and Xl-Oct-1B synthetic transcripts (produced from the original excised clones in pBluescript sk- with T7 or T3 RNA polymerase). These assays are shown in figure 20A. Sense Xl-Oct-1A synthetic RNA protects a major band of about the predicted size (265 bases), which is smaller than the probe. There is no signal with the 'no RNA' control assays. Sense Xl-Oct-1B synthetic RNA protects a smaller fragment (about 50 bases). The source of this protected fragment is not obvious from a comparison of the Xl-Oct-1A and Xl-Oct-1B sequences, since

Figure 20.

Detection of Xl-Oct-1A and Xl-Oct-1B transcripts in oocytes by RNase protection assays.

- (A) RNase protection assays with Xl-Oct-1B and Xl-Oct-1A synthetic sense transcripts. 50pg of synthetic RNA was hybridised with 250 000 cpm of internally labelled antisense synthetic RNA probe (265 bases of Xl-Oct-1A + about 95 bases of vector polylinker sequence) and then digested with RNases (A and T1). The reaction products were electrophoresed on an 8% denaturing polyacrylamide gel. An autoradiograph of the gel is shown. Duplicate assays were carried out. Undigested probe and mock hybridisation with no RNA controls are included. The positions of fragments protected by Xl-Oct-1A and Xl-Oct-1B transcripts is indicated. Markers are pBR322 digested with Hpa II and labelled by end-filling.
- (B) RNase protection assays with a mixture of Xl-Oct-1A and Xl-Oct-1B synthetic sense transcripts (50pg each RNA) and total oocyte RNA (20ug, equivalent to 4 oocytes). Assays were carried out as described above. The positions of protected fragments from Xl-Oct-1A and Xl-Oct-1B transcripts is indicated. Undigested probe and mock hybridisation with no RNA controls are included.

A**B**

Results and Discussion

they are not perfectly matched over a region as long as 50 bases. Presumably the RNases do not cut, or cut less efficiently at particular short mismatches between probe (X1-Oct-1A) and transcript (X1-Oct-1B). The test shows, therefore, that this probe should be able to be used to detect and distinguish X1-Oct-1A and X1-Oct-1B transcripts.

8.3 X1-Oct-1A and X1-Oct-1B transcripts can be detected in X. laevis oocytes.

Assays were set-up with a mixture of X1-Oct-1A and X1-Oct-1B synthetic sense transcripts (as control) and with total oocyte RNA. The result is shown in figure 20B. The mixture of synthetic transcripts and oocyte RNA give similar protection patterns, indicating that transcripts of both clones are present in, and derived from X. laevis oocytes. The synthetic transcript assay contains approximately 50pg of each RNA species. Consequently, by a comparison of signal intensity, the 20ug of total oocyte RNA contains roughly 10pg of each Oct-1 transcript, which is equivalent to 2.5pg per oocyte. Since an oocyte contains about 50ng (1% of 5ug total RNA) of A+ RNA the abundance of each Oct-1 transcript is roughly 1 in 20 000. It is possible that the transcripts are present solely in the layer of follicle cells which surround the oocyte (these were not removed before preparing the RNA used for cDNA synthesis). However, this is unlikely since follicle cell RNA constitutes less than 1% of oocyte RNA, and if transcripts were present solely in follicle cells their

Results and Discussion

abundance would be 1 in 200 transcripts. It also seems improbable that Oct-1 transcripts are present solely in the follicle cells since, as will be shown later, there are roughly equivalent amounts of Oct-1 protein in the oocyte and unfertilised egg (which has lost the layer of follicle cells).

Results and Discussion

Chapter 9.

Expression of Xl-Oct-1A synthetic RNA in micro-injected oocytes.

To confirm that Xl-Oct-1A contained an open reading frame of the predicted length it was decided to attempt to express synthetic sense transcripts in micro-injected oocytes.

The Xl-Oct-1A cDNA in the vector pBluescript sk- (Stratagene) was linearised with Sac I, and transcribed with T7 RNA polymerase to produce sense transcripts. Approximately 10nl of capped transcript (concentration approximately 25ng/ul) was micro-injected into the cytoplasm of X. laevis oocytes, and translation products labelled by overnight incubation in medium containing ^{35}S -methionine. Labelled proteins were resolved by SDS-PAGE and detected by fluorography. The result is shown in figure 21B. There was no detectable translation of synthetic transcripts, since an extra band cannot be seen with protein extract from injected oocytes, relative to protein extract from un-injected oocytes.

Previous experience in our laboratory (R. W. Old and G. Sweeney, pers. comm.) has shown that shortening of the 5' untranslated leader of clones to just before the initiation codon can enhance expression of synthetic transcripts in micro-injected oocytes. Consequently it was decided to shorten the leader of Xl-Oct-1A in an attempt to enhance expression.

The Xl-Oct-1A cDNA (flanked by Eco RI sites) was subcloned

Figure 21.

Expression of X1-Oct-1A synthetic transcripts in micro-injected oocytes.

- (A) X1-Oct-1A was subcloned into the Eco RI site of M13mp18 and site directed mutagenesis used to create the Sal I site indicated. This site was used to shorten the leader sequence of the clone, removing an upstream out of frame ATG codon.
- (B) X1-Oct-1A in pBluescript SK- (stratagene) was linearised with Sac I and transcribed with T7 RNA polymerase to generate full-length transcripts.

A fragment from the Sal I site made in the leader of X1-Oct-1A to the Sal I site in the M13mp18 polylinker was subcloned into pBluescript KS+ (Stratagene). This subclone was linearised with Kpn I and transcribed with T7 RNA polymerase to generate shortened leader transcripts.

Approximately 10nl of synthetic capped transcript (RNA concentration approximately 25ng/ul) was micro-injected into the cytoplasm of *X. laevis* oocytes. Oocytes (injected and not injected controls) were incubated overnight in medium containing ^{35}S -methionine, protein extract was prepared from pooled oocytes and the proteins resolved by SDS-PAGE (10% separating gel). Labelled proteins were detected by fluorography. The size and position of ^{14}C methylated protein markers is indicated. * indicates the position of an extra labelled protein band present in extract from oocytes which had been injected with shortened leader transcripts.

A

Out of frame ATG.
Unfavourable initiation context.

GCAGCGGCGCCAGCAGGCACGCAGATCTAAAAAGGAGCAGGTTTGGTGCACCAAATCAT
stop

TTTCTAAGTCTATCCAGCGTTAATTTTTTACCAGGAAGAGCCTGAAATAAGTAGAACACTC

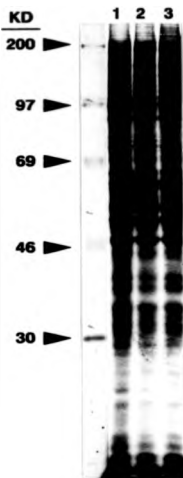
TGGGAAGAGCAGTCGTCCTTAATACACCATGAAATTCAT.....
stop M K L H.....

mutagenesis

In frame ATG.
Favourable initiation context.

A = SalI site used to shorten leader.

B



* 79 KD
PRODUCT

1=not injected

2=shortened leader
transcript injected

3=full-length
transcript injected

Results and Discussion

into M13mp18, and site directed mutagenesis used to convert the sequence GTCGTC (16 bases upstream of the first in frame ATG) to a Sal I site (GTCGAC). This is illustrated in figure 21A. A fragment from this Sal I site to the Sal I site in the M13mp18 polylinker was subcloned into pBluescript ks+ (Stratagene). Sense transcripts were produced with T7 RNA polymerase from DNA template which had been linearised with Kpn I. Capped transcript, with the shortened leader, was micro-injected into the cytoplasm of X. laevis oocytes at an equivalent concentration to the full-length transcript. Translation products were detected as described above. An extra protein band (relative to un-injected oocyte extract, and extract from oocytes injected with the full-length transcript) in extract from oocytes injected with the shortened leader transcript indicates that this transcript is expressed in oocytes. The size of the product exactly corresponds to that predicted from the open reading frame determined by sequencing (79kD).

The reason why the full-length transcript is expressed at lower levels than the shortened leader transcript is unclear. There is an upstream out of frame AUG codon (see figure 21A), but this would not be expected to inhibit translation (Kozack (1989)). Kozack suggests that the 40s ribosomal subunit binds to the 5' end of mRNA and scans along until it finds the first AUG codon in a favourable context for the initiation of translation. A favourable context for initiation is determined by the sequence surrounding the AUG. The most

Results and Discussion

important residues are a purine three bases upstream of the AUG and a G immediately after the AUG (PuXXAUGG). The purine (usually A) upstream of the AUG is most important, with other residues having little effect on the efficiency of translation if this residue is present. However, in the absence of the purine the downstream G is essential. Consequently, the upstream AUG is in an unfavourable context for the initiation of translation, and 'leaky' scanning should allow the ribosome to reach the correct AUG. 'Leaky scanning' is also more efficient when, as is the case here, the AUG is followed by a stop codon. The first in frame AUG is in a favourable context for the initiation of translation, having an A in the correct position upstream of the AUG. Consequently there is no obvious reason why the full-length transcript is expressed at a lower level than the transcript with the shortened leader. Inefficiency of translation is not restricted to the micro-injected oocytes. Neither synthetic transcript (full-length or shortened leader) can be expressed in vitro in rabbit reticulocyte lysate (not shown).

Results and Discussion

Chapter 10.

Preparation of Xl-Oct-1A fusion protein constructs.

Introduction.

It was decided to express Xl-Oct-1A fusion proteins in E. coli for two reasons. Firstly, in order to produce large amounts of Xl-Oct-1A for use as an antigen. The antigen fusion protein would be injected into rabbits to raise a polyclonal antiserum. At the time of preparing these fusion proteins human Oct-1 and Oct-2 had been cloned and the highly conserved POU domain recognised. Consequently, a fusion protein lacking the POU domain would be required if the antiserum was to be specific to Oct-1. Secondly, I wanted to show that Xl-Oct-1A encoded a functional octamer-binding protein. Fusion proteins (containing the POU domain) would be used to demonstrate this.

This chapter describes the fusion protein constructs made. The use of fusion proteins to raise a polyclonal antiserum, and in band shift assays will be described in later chapters.

Most fusion proteins made were T7 gene 10 fusions, made in the high level expression pET-3 series vectors (Rosenberg et al (1987)). The pET-3 series vectors (a, b, c) have a Bam HI cloning site in each of the reading frames. Fusion proteins are produced from clones in these vectors in the E. coli strain BL21 (DE3) (Studier and Moffatt (1986)). BL21 (DE3) has an integrated copy of T7 RNA polymerase under the control of the lac promoter. IPTG induces expression of T7 RNA polymerase from the lac promoter, which transcribes the

Results and Discussion

message for the fusion protein from the T7 gene 10 promoter in pET-3 series vectors. A beta-galactosidase fusion protein was also made in the vector pBluescript sk- (Stratagene). It was intended to show that the polyclonal antiserum raised reacted with Oct-1 by inhibiting the formation of octamer element:fusion protein complexes (in band shift assays). To demonstrate that the antiserum was not reacting to the T7 gene 10 part of the fusion protein, this beta-galactosidase fusion protein was also made.

10.1 Making a Bgl II site at the 5' end of X1-Oct-1A.

The pET-3 series vectors have a Bam HI cloning site. Consequently, to produce X1-Oct-1A fusion proteins extending from the 5' end of the coding sequence, it was necessary to introduce a Bam HI compatible site by oligonucleotide directed mutagenesis. A Bgl II site was selected, rather than a Bam HI site, for reasons which will become apparent later. The Bgl II site made allowed fusion proteins to be made by cloning into the Bam HI site of pET-3B, and by cloning into the Bam HI site of pBluescript sk- (Stratagene). The X1-Oct-1A cDNA was subcloned into the Eco RI site of M13mp18 and oligonucleotide directed mutagenesis used to convert the sequence AGATGC at nucleotide 196 (15 amino acids downstream of the first in frame methionine) to a Bgl II site (AGATCT). This is illustrated in figure 22.

Figure 22.

Xl-Oct-1A sequence contained in fusion protein constructs.

The amino acid sequence of the Xl-Oct-1A open reading frame is shown.

The generation of a Bgl II site by oligonucleotide directed mutagenesis is illustrated.

The 5' and 3' ends of Xl-Oct-1A coding sequence contained in fusion protein constructs (pET-1 to 4 and pBS-1) is indicated by arrows. The numbers adjacent to the arrows indicate the number of amino acids from the start of the open reading frame. pET-1, 2 and 3 were made in the vector pET-3B (Rosenberg *et al* (1987)). pET-4 was made in the vector pET-3A (Rosenberg *et al* (1987)). pBS-1 was made in the vector pBluescript SK⁻ (Stratagene). The location of the POU domain is shown.

TCAGATCTA Bgl II pET-1,2,3 and pBS-1 5' and
mutagenesis

TCAGATCTA

001 MELHSSSLIQNHAWLSDARHNNPSETSKSPESGDCGNTGTQINCLDPQKQAVPIGAITSAAQALLGHLHQ
071 VQLAGTSLQAAAHSLNVQTEKEREFGPEPMQVVQPSQQFSLQAAIPQTQLMVAGGQIAGLTLTTPAQQMLL
141 QQAQALLAAAVQHSANQHHNAAGATISASAATPMTQIFLSQPIQIAQDLQQLQQLQQNLNLQYVLVH
pET-3 3' end Stu I(231)
211 PITNLQSAQFIISQTPQGGQLLQAQMLLTQLPQQSQANLLQSQPSITLTSQPAITPTTIAATPVQQLPQ

281 SQTTPKRIDTFSLEFSDLELEQFARYPKQAIKLGFTQGDVGLANGKLYGNDFSQTIISFEALWLSF
FOU DOMAIN
351 KNNCKLEPLAKKVLNDVLENNIYSSTLTQSVINSPCRNGECLNTERKKETSTETNINVALNKSFLKNG

pET-2 3' end and pET-4 5' end Bam HI(463)
421 KPTSENIHMIADQLNKKKIVIVVVCNRRQKKRIINFFSSGGSSSSPIKSLFSSPNPLVASTFSLVSSP
491 ATILTVPNVLPLTISAAAITSPHPIGTTGTSSANTATVISTAPPVSSVLTSPLSSSPSATAASSEASTAG
561 ETSTTQTITSTPMTSSLNTGQVNVATSGINTAAATALQGAALPTNASLAAMAAAAGLNPLGNAPSQFAAG
631 GALTSLNPGALGALSALPMSNSTLATIQALASSGSLPITSLEDAAGNLVFANAGCTPHIVTAPLFLNPQM

pET-1,4 and pBS-1 3' end(760)
701 LSLFTSNFVSLISAASAGATGPIITSLMATTSSIDSIGNALPTMASASGAASTTTSASKAQ

Results and Discussion

10.2 Making fusion protein constructs.

The fusion protein constructs made are illustrated in figure 22. A fragment from the Bgl II site made by mutagenesis to the Bam HI site in the Xl-Oct-1A coding sequence (just 3' to the POU domain) was subcloned into the Bam HI site of pET-3B to create pET-2. The pET-2 fusion protein contains amino acids 16 to 463 of Xl-Oct-1A fused to the first eleven amino acids of T7 gene 10. pET-2 can be cut with Bam HI to linearise at the 3' end of the cDNA insert (there is no Bam HI site at the 5' end as a result of the ligation of Bam HI and Bgl II sticky ends). This allows deletion from the 3' end and addition of the remaining 5' end of the Xl-Oct-1A coding sequence as described below.

pET-2 was cut with Bam HI (at the 3' end of the cDNA insert) and Stu I (within the Xl-Oct-1A coding sequence contained in the clone), the Bam HI end blunted with the Klenow fragment of DNA polymerase I, and the ends ligated to create pET-3. pET-3 contains amino acids 16 to 231 of Xl-Oct-1A.

A fragment was isolated from the Xl-Oct-1A cDNA clone in the vector pBluescript SK-. The fragment extended from the Bam HI site in the Xl-Oct-1A coding sequence to the Bam HI site at the 3' end of the cDNA insert in the pBluescript SK-polylinker. pET-2 was cut with Bam HI (at the 3' end of the cDNA insert) and this fragment inserted to create pET-1. pET-1 contains amino acids 16 to 760 (3' end) of Xl-Oct-1A and the 3' untranslated region.

The fragment isolated to insert into pET-2 to create pET-1

Results and Discussion

was inserted into the Bam HI site of pET-3A to create pET-4. pET-4 contains amino acids 463 to 760 (3' end) of X1-Oct-1A and the 3' untranslated region.

A fragment from the Bgl II site of the mutated X1-Oct-1A cDNA in M13mp18 to the Eco RI site at the 3' end of the cDNA insert was subcloned into pBluescript SK- to create pBS-1. pBS-1 contains amino acids 16 to 760 (3' end) of X1-Oct-1A, and the 3' untranslated region, in frame to the beta-galactosidase gene of pBluescript SK-.

Results and Discussion

Chapter 11.

Production of an anti Xl-Oct-1A polyclonal antiserum.

Introduction.

An anti Xl-Oct-1A antiserum was primarily required to identify the Oct-1 band in band shift assays, and also to stain Western blots of protein extract. To exclude the possibility of the antiserum reacting with other proteins via a conserved POU domain, the antiserum was raised against an Xl-Oct-1A fusion protein lacking the POU domain.

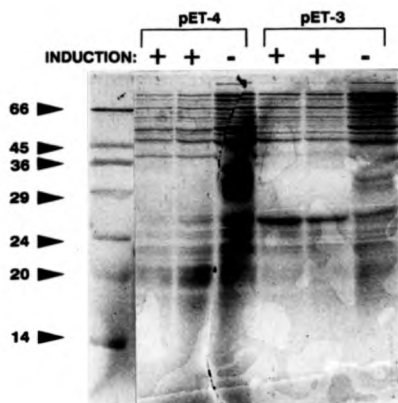
11.1 Making an Xl-Oct-1A fusion protein for use as an antigen.

Two fusion protein constructs (see chapter 10) were suitable for use as antigens (ie. they lack the POU domain): pET-3 and pET-4. pET-3 contains amino acids 16 to 231 of Xl-Oct-1A in the vector pET-3B (Rosenberg et al (1987)), and pET-4 amino acids 463 to 760 in the vector pET-3A (Rosenberg et al (1987)). These constructs were transformed into the E. coli strain BL21 (DE3) (see chapter 10). Induced cultures (IPTG added to induce expression of T7 RNA polymerase from the lac promoter, which then synthesises message for the T7 gene 10 fusion protein) and un-induced cultures were grown, and proteins analysed by SDS-PAGE. A photograph of the gel stained with Coomassie blue is shown in figure 23. The predicted size of the pET-3 fusion protein (215 amino acids of Xl-Oct-1A and 11 amino acids of T7 gene 10) is approximately 25kD (assuming the average molecular weight of

Figure 23.

Expression of fusion proteins suitable for use as antigens.

Three cultures of BL21(DE3) carrying pET-3 or pET-4 were grown to $A_{600} = 0.8$. IPTG was added to two cultures carrying each construct to a final concentration of 1mM (+). One of the three cultures was not induced (-). Growth was continued for a further two hours, after which time 100ul of culture was microfuged and suspended in SDS gel loading buffer, and electrophoresed on a 15% separating SDS-polyacrylamide gel. A photograph of the Coomassie stained gel is shown. Size and position of protein molecular weight markers is shown.



Results and Discussion

an amino acid is 110). A more intense band of this size is seen in the induced tracks relative to the un-induced track, indicating high level expression of the fusion protein. The predicted size of the pET-4 fusion protein (297 amino acids of Xl-Oct-1A and 11 amino acids of T7 gene 10) is approximately 34kD. There is no visible induced band of this size, indicating that this fusion protein is not strongly expressed. This could be a consequence of instability (of message or protein), or due to toxicity of the fusion protein. Toxic proteins are known to select for the loss of the plasmid expressing the fusion protein, or to select for promoter mutations which reduce expression if, as was the case here, the ampicillin selection is maintained at high levels (Studier and Moffatt (1986)). The presence of the pET-4 plasmid caused a much reduced growth rate, which is consistent with the fusion protein being toxic. Consequently the pET-3 fusion protein was selected for use as an antigen.

A 50ml culture of BL21 (DE3) carrying pET-4 was grown to $A_{600} = 0.8$, and then induced by the addition of IPTG (to a final concentration of 1mM). Proteins were resolved on an SDS-polyacrylamide gel, and the gel lightly stained with Coomassie blue in water. A gel slice containing the fusion protein was cut out. A small fraction of the culture was run alongside a known amount of marker protein to estimate the amount of fusion protein contained in the gel slice. The fusion protein, isolated in this way, was used as an antigen.

Results and Discussion

11.2 Production of a polyclonal antiserum.

pET-3 fusion protein (in the gel slice) was emulsified with Freund's adjuvant as described in Methods. A New Zealand White rabbit was given three injections of 50 to 100ug of protein at two week intervals. For the first injection protein was emulsified with complete adjuvant, and for subsequent injection with incomplete adjuvant. Prior to the first injection approximately 10mls of pre-immune serum was taken. Two weeks after the final injection immune serum was taken.

Antiserum was checked by using it to stain a Western blot of protein extract from induced and un-induced E. coli BL21 (DE3) carrying the pET-3 fusion protein construct. Proteins were resolved by SDS-PAGE (15% separating gel) and electrophoretically transferred to a nitrocellulose filter. The filter was incubated with antiserum, and bound antibodies detected by incubating with a biotinylated secondary antibody, then streptavidin alkaline phosphatase conjugate, and then colour development with BCIP and NBT (which produce a coloured deposit on the filter in a reaction catalysed by alkaline phosphatase). A photograph of the stained filter is shown in figure 24. The antibody reacts with a band of the correct size (the pET-3B fusion protein is 25kD) in both induced and un-induced culture extracts. The reaction is stronger with induced culture extract, indicating that there is a lower basal level of fusion protein expression in the absence of induction with IPTG (not detected by Coomassie

Figure 24.

The anti Oct-1 antiserum detects the E. coli fusion protein against which it was raised.

E. coli BL21 (DE3) cultures carrying pET-3 were microfuged, resuspended in 0.01 volumes of SDS loading buffer and the amount indicated resolved on a 15% separating SDS-polyacrylamide gel. Induction indicates that IPTG (to a final concentration of 1mM) was added to the culture at $A_{600} = 0.8$. Growth of induced and not induced cultures was continued for two hours after this point. The gel was electrophoretically transferred to a nitrocellulose filter, and the filter incubated with a 1 in 100 dilution of the anti Oct-1 antiserum. Bound antibodies were detected as follows: incubation with biotinylated anti rabbit immunoglobulin; incubation with streptavidin alkaline phosphatase conjugate; and finally colour development with NBT and BCIP (which give coloured products in a reaction catalysed by alkaline phosphatase). Size and position of protein molecular weight markers is shown.

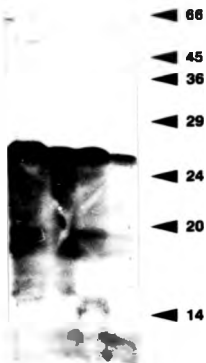
INDUCTION:

+

-

SAMPLE VOLUME:
(MICROLITRES)

5 1 5 1



Results and Discussion

staining). Weaker reaction to proteins below the strong 25kD band is probably to degraded fusion protein. Consequently the antiserum reacts with Xl-Oct-1A produced in E. coli.

Antiserum was purified by affinity for the pET-3 fusion protein electrophoretically transferred to a nitrocellulose filter as described in Methods. Purified antibody was routinely used to stain Western blots of X. laevis protein extract in an attempt to ensure that the bands detected were related to Oct-1, and were not being detected by antibodies present in the antiserum before immunisation. However, no significant difference was observed between the proteins detected by whole and purified antiserum (see later).

Results and Discussion

Chapter 12

The *Xenopus laevis* Oct-1 cDNA encodes a functional octamer-binding protein.

Extract from *E. coli* expressing the fusion protein constructs pET-1, pET-2 and pBS-1 (described in chapter 10) were used in band shift assays to show that X1-Oct-1A encodes a functional octamer-binding protein. The Oct-1 sequence contained in these constructs is illustrated in figure 25A. pET-1 is the full-length cDNA in pET-3B (actually it lacks the first 16 amino acids from the N-terminal end), and pBS-1 is the full-length cDNA in pBluescript SK-. pET-2 and pET-3 are equivalent to pET-1 which has been progressively deleted from the C-terminal end. pET-3 does not contain the DNA-binding POU-domain, and is therefore included as a control. Extract from *E. coli* expressing pET-1, pET-2 and pBS-1 binds to an octamer probe (end-labelled 25mer duplex oligonucleotide containing the consensus octamer, sequence given later in figure 28F), and as expected pET-2 extract gives a greater mobility shifted band, since the Oct-1 sequence contained in this construct is deleted at the C-terminal end (figure 25B). This binding is specific as it can be competed-out by inclusion of an excess of unlabelled probe in the binding reaction. As controls, extract from the host *E. coli* strain (TG2) of pBS-1 and from *E. coli* expressing pET-3 (which does not contain the DNA binding POU-Domain) do not bind to the octamer probe in band shift assays.

Figure 25

Oct-1 sequence contained in fusion protein constructs and their binding to a probe containing an octamer motif analysed by band shift assays.

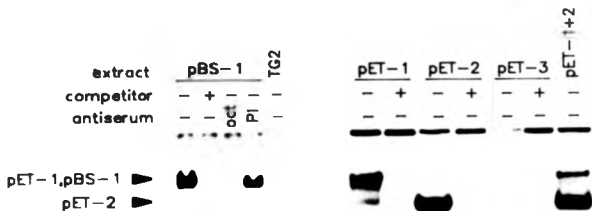
- (A) Oct-1 fusion proteins contain Oct-1 sequence extending from the Bgl II site made by oligonucleotide directed mutagenesis to the position on the Xenopus laevis Oct-1 amino acid sequence indicated by the name of the construct (pET-1, pET-2, pET-3 and pBS-1). The location of the FOU (DNA binding) domain is shown.
- (B) Band shift assays with extract from E. coli expressing the fusion protein constructs indicated, and a probe containing an octamer motif (oct probe, see figure 28F). Competitor (an excess of unlabelled probe) and antiserum (Oct = anti Xenopus laevis Oct-1 (raised against the pET-3 fusion protein); PI = pre-immune serum) were included in the binding reaction where indicated. TG2 extract is from the host E. coli strain of pBS-1. The positions of undegraded pBS-1, pET-1 and pET-2 complexes and free probe (F) are indicated.

A

TCAGATGA Bgl II
 || mutagenesis
 TCAGATGCA

001 MKLHSSSLIQNHAWLSDARHNNPSETSKSPESGDGNTGTQTNGLDFOKQAVFIGAITSAAQAQLLGHLLHQ
 071 VOLACTSLQAAASLNVQTKPKKPGEPHQVQPSQPSLQAAATPQQLNVAGGQIAGLTITPAQQLML
 141 QQAQAQLLAAAVQHSARQHNAAAGATISASAATPMTQIPLSQPIQIAQDLQQLQQLQQNNLNLQQYVLVH
 211 PTINLQSAQV IISQTPGQGGLLQAQMLLTQLPQSSQAMLLQSQPSIILTSQPATPTETIAATPVQQLPQ
 281 SQITPKRIDTPSLWPSDLRLRLQVAKTPQQRITLGLFTQGDVGLANGKLYGNDPSQYIYSEPEALNLSF
 351 YNNCKELKPLLEKWLMDAVLENIITSDSTLTWQSVLMSYCHGNGECLNENKKEYSYETNIEVALKESVLENGQ
 421 EPTSEETIYIADQLNNKKEVINWVPCNKKKRNWPFSSGGSSSSPIKSLPSSPNPLVASTPSLVTSSSP
 491 ATTLTTHVPLPLTISAAATSPHPTGTGTSSANTATVISTAFFVGSVLTSPSLSSSPATASSEASTAG
 561 ETSITQTSTPMTSSSLWTGQVMVTASGIRTAATATLQGAALPTNASLAAMAAAAGLWPGLNAPSQPAAG
 631 GALFSLNPGALGSALPMSNSTLATIQLASSGSLPITSLEAAGNLVFANAGGTFNIVTAPLPLMPQN
 701 LSLPTSNPVSLISAASAGATGPITSLMATTSSIDSIQNALFTMASASGAASSTTSASKAQ

B



F

Results and Discussion

The assays shown here use extract from un-induced E. coli cultures. Induction with IPTG (which induces pBS-1 expression directly, and pET construct expression indirectly by inducing T7 RNA polymerase) was found to have either no effect on the yield of these fusion proteins, or to give lower yields/degraded products. This is probably because the proteins produced are toxic to E. coli. This leads to the loss of the expressing plasmid, or if the ampicillin selection is maintained at high levels cells carrying plasmids with promoter mutations which reduce expression are selected (Studier and Moffatt (1986)). In the case of degraded products the toxic proteins had probably caused cell death.

The antiserum raised against the pET-3 fusion protein included in the pBS-1 band shift assay inhibits formation of the Oct-1 - DNA complex, whereas pre-immune serum does not. This demonstrates that the antibody reacts to XI-Oct-1A, and not to the non-Oct-1 part of the fusion protein, since the antibody was raised against a T7 gene 10 fusion protein, and reacts with the pBS-1 fusion protein which is a beta-galactosidase fusion protein.

Functional expression of Oct-1 in E. coli suggests that no other factors are required for DNA binding activity. Mixing of pET-1 and pET-2 extracts in band shift assays gives two strong shifted bands of the mobility seen in un-mixed extracts (figure 25B). This shows that Oct-1 binds to the probe as a monomer, since if Oct-1 bound as a dimer a third

Results and Discussion

band of intermediate size corresponding to a pET-1:pET-2 heterodimer would be expected. The fact that pET-2 binds and pET-3 does not shows that the Xl-Oct-1A binding activity is located in a region of the protein containing the POU domain (which has been shown to be necessary and sufficient for the binding of human Oct-1, see Sturm et al (1988A) and Sturm and Herr (1988)).

Results and Discussion

Chapter 13.

Two octamer-binding proteins can be detected in *Xenopus laevis* oocyte extract.

13.1 Oocytes contain Oct-1 and a second octamer-binding protein.

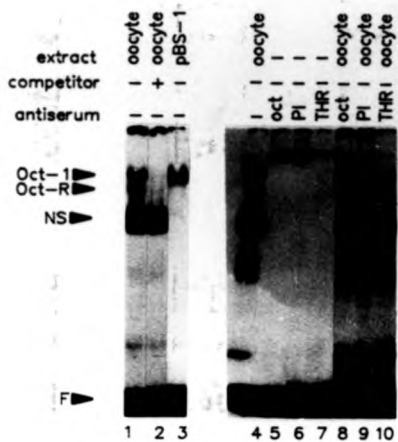
Band shift assays with a probe containing the consensus octamer (oct probe, see figure 28F) and oocyte protein extract detect two specific bands as a result of octamer binding proteins. Specific binding can be competed-out by inclusion of an excess of unlabelled probe in the binding reaction (figure 26). One strong and several weak non-specific binding activities (of greater mobility than the two specific binding activities) are also detected.

The upper specific band is of a similar mobility to the pBS-1 fusion protein band. The pBS-1 fusion protein is of similar size to the protein encoded by the X1-Oct-1 cDNA (in pBS-1 16 amino acids of X1-Oct-1 coding sequence are replaced by 31 amino acids of the beta-galactosidase gene of pBluescript SK-, resulting in a protein of 775 residues rather than 760). Formation of the upper band (and not the lower band or the non-specific binding) is inhibited by inclusion of antiserum raised against the pET-3 Oct-1 fusion protein in the binding reaction. Pre-immune serum and an antiserum raised against *Xenopus laevis* thyroid hormone receptor (gift of R. W. Old) do not inhibit formation of this complex. An extra low mobility band is seen in assays

Figure 26

Band shift assays showing that oocyte extract contains Oct-1 and a second octamer-binding protein, Oct-R.

The probe contains an octamer motif (oct probe, see figure 28F) and extract is from *Xenopus laevis* oocytes, except pBS-1 extract (from *E. coli* expressing the pBS-1 fusion protein). Assays in lanes 5-10 contain antiserum (PI = anti *Xenopus laevis* Oct-1 pre-immune serum; oct = anti *Xenopus laevis* Oct-1 (raised against the pEI-3 fusion protein); THN = anti *Xenopus laevis* thyroid hormone receptor) and in lanes 5-7 contain no protein extract to show that the sera bind non-specifically. Lane 2 contains an excess of unlabelled probe in the binding reaction as competitor. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



Results and Discussion

containing antiserum. Assays with serum and no extract demonstrate that this is due to non-specific binding of a protein in the serum.

Consequently the upper band seen in these band shift assays is identified to be a result of binding of Xenopus laevis Oct-1. The Oct-1 band probably represents binding of the products of both X1-Oct-1A and X1-Oct-1B clones. The second greater mobility band is as a result of binding of a previously unidentified octamer-binding factor, which is antigenically distinct from the N-terminal domain of Oct-1 contained in the pET-3 fusion protein. This factor has been termed Oct-R (for octamer-related).

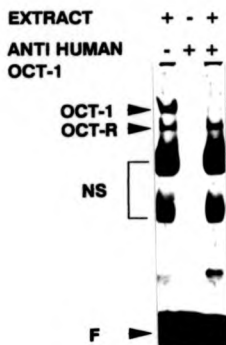
13.2 An antibody against full-length human Oct-1 recognises Xenopus laevis Oct-1, but not Oct-R.

The effect of an antiserum raised against a full-length human Oct-1 fusion protein (gift of W. Herr, see Sturm et al (1988)) included in band shift assays with the octamer probe (oct probe, see figure 28F) and oocyte protein extract is shown in figure 27. As with anti Xenopus laevis Oct-1 antiserum, the serum alone binds non-specifically to the probe giving a low mobility retarded band. The anti human Oct-1 antiserum inhibits the formation of the Oct-1 complex, but not the Oct-R complex. This confirms the identification of the upper specific band as being the result of Oct-1 binding. The anti human Oct-1 antiserum was raised against a full-length Oct-1 fusion protein, including the POU-domain.

Figure 27.

Anti human Oct-1 antiserum reacts with Xenopus laevis Oct-1, but not Oct-R.

Band shift assays with an octamer probe (oct probe, see figure 28F) and with oocyte extract and/or anti human Oct-1 antiserum included where indicated. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



Results and Discussion

This antibody inhibits the formation of both human Oct-1 and Oct-2A complexes (Sturm et al (1988A)), presumably as a result of the reaction of antibodies against the conserved POU-domain. It is possible that Oct-R binds to DNA via a POU-domain, however the anti human Oct-1 antiserum does not react with Oct-R. This result may indicate that the Xenopus laevis Oct-R POU domain (unlike the Xenopus laevis Oct-1 POU domain) is significantly diverged from the human Oct-1/Oct-2 POU domain.

Results and Discussion

Chapter 14.

A comparison of the binding properties and distribution of Oct-1 and Oct-R.

14.1 Oct-1 and Oct-R have different binding affinities.

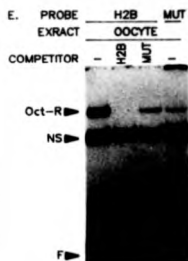
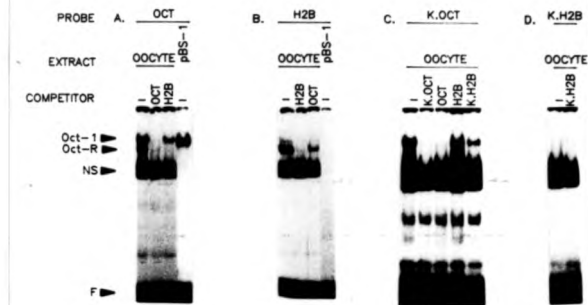
In Xenopus laevis histone H2B gene promoters the octamer motif occurs as the core of an extended consensus sequence known as the H2B box (see introduction section 1.1.2, and figure 33). The most common octamer motif in Xenopus laevis H2B boxes is not the perfect octamer (ATTGCGAT) but a 7 out of 8 match (GTTGCGAT). I have tested the binding of Oct-1 and Oct-R to these octamer motifs in the context of the Xenopus laevis H2B box (oct probe = ATTGCGAT, H2B probe = GTTGCGAT in this context) and in the context of the mouse immunoglobulin heavy chain enhancer (k.oct probe = ATTGCGAT, k.H2B probe = GTTGCGAT in this context). The results are shown in figure 28. Equal amounts of the different probes, and an equal amount of oocyte extract was used in assays where comparisons are made.

As shown previously, the oct probe binds both Oct-1 and Oct-R. The H2B oligo is a poor competitor for Oct-1 binding, but an efficient competitor for Oct-R binding (figure 28A). Correspondingly Oct-R binds well to the H2B probe and Oct-1 does not detectably bind. Oct-R binding to the H2B probe is stronger than to the oct probe (the gel exposures in figure 28 A and B are directly comparable). The oct-1 fusion protein, pBS-1, does not detectably bind to the H2B probe (figure 28B). With the H2B probe a faint, variable and not

Figure 28.

Comparison of the binding affinities of Oct-1 and Oct-R for different octamer-containing oligos using band shift assays.

Oct and k.oct oligos contain the perfect octamer in the context of the histone H2B gene promoter and the immunoglobulin heavy chain enhancer, respectively. H2B and k.H2B oligos contain a 7 out of 8 match octamer motif in the context of the histone H2B gene promoter and the immunoglobulin heavy chain enhancer, respectively. The mut oligo is the same sequence as the H2B oligo except with two point mutations introduced into the octamer motif. Protein extract is from *Xenopus laevis* oocytes, except pBS-1 extract (from *E. coli* expressing the pBS-1 fusion protein). A to E show band shifts using the oligos indicated as probe, and with the unlabelled oligos indicated as competitor. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated. F summarises the binding affinities of Oct-1 and Oct-R to the oligos.



F. PROBE

PROBE	SEQUENCE	OCT-1 BINDING	OCT-R BINDING
H2B	GAG CTGCCTC GTTGCAT GGA GGA	+	++++
OCT	GAG CTGCCTC GTTGCAT GGA GGA	++++	++
K-H2B	CAC CTGGTA GTTGCAT T TCT AAA	+	-
K-OCT	CAC CTGGTA GTTGCAT T TCT AAA	+++	-
MUT	GAG CTGCCTC GTTGCAT GGA GGA	-	+

OCTAMER
H2B-BOX

Results and Discussion

efficiently competable band is seen above Oct-R and just below the Oct-1 position. This binding is not inhibited by the anti Oct-1 antiserum (not shown). The k.oct probe detectably binds only Oct-1. As expected the oct oligo is an efficient competitor for this binding and the H2B and k.H2B oligos are poor competitors. Oct-1 binds slightly less efficiently to the k.oct probe than the oct probe (figure 28C is a long exposure to illustrate the different competitive abilities of the cold oligos. With an equivalent exposure of A and C it is apparent that Oct-1 binds slightly more efficiently to the oct probe than to the k.oct probe). The k.H2B probe does not detectably bind Oct-1 or Oct-R (figure 28D).

To show that Oct-R binding is octamer-dependent the H2B oligo with two point mutations in the octamer motif (mut probe, GTTTGCAT > CGTTGAAT) was used as probe and competitor in band shift assays. This oligo binds Oct-R less efficiently than the H2B oligo and is a poor competitor for Oct-R binding (figure 28E).

The binding affinities of Oct-1 and Oct-R for the different oligos are summarised in figure 28F. Oct-1 binds efficiently to the perfect octamer in either sequence context. Binding is slightly more efficient in the context of the H2B box. Oct-1 does not bind to the 7 out of 8 octamer most often seen in Xenopus laevis histone H2B gene promoters. Oct-R binds efficiently to both the perfect and 7/8 octamer motifs, but only binds in the context of the H2B gene

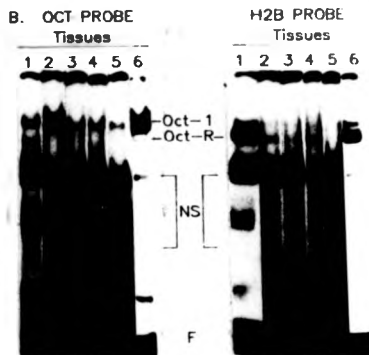
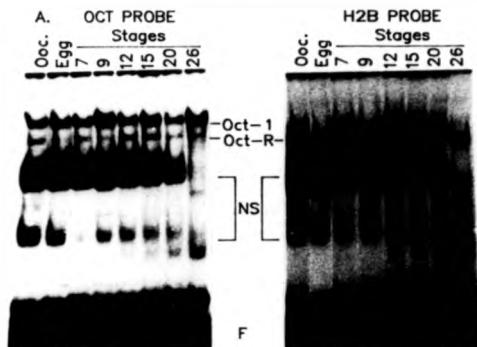
Figure 29.

Distributions of Oct-1 and Oct-R determined by band shift assays.

Oct and H2B probes (see figure 28F) were used in band shift assays to maximise the binding of Oct-1 and Oct-R, respectively.

- (A) With extract from Xenopus laevis oocytes, eggs and stages of early development (Nieuwkoop and Faber (1956)). The equivalent of one oocyte/egg/embryo was used in each assay.
- (B) With extract from adult Xenopus laevis tissues and the cell line, Xtc.

The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



TISSUES: (1) OOCYTE
(2) OVIDUCT
(3) KIDNEY
(4) BRAIN
(5) BLOOD
(6) Xtc CELLS

Results and Discussion

promoter. Binding is slightly more efficient to the 7/8 octamer.

14.2 Distribution of Oct-1 and Oct-R in tissues and early development.

Band shift assays have been used to analyse the distribution of Oct-1 and Oct-R in early Xenopus laevis development and in adult Xenopus laevis tissues. Oct and H2B probes (figure 28F) were used to maximise the binding of Oct-1 and Oct-R, respectively. Both Oct-1 and Oct-R are present at an approximately constant level throughout early development (see figure 29A, the assays shown contained extract from the equivalent of one oocyte/embryo).^{*INSERT.} Oct-1 is present in all adult tissues tested (oocyte, oviduct, kidney, brain, blood) and in the cell line, Xtc. Oct-R has a similar distribution, but cannot be detected in the extract from blood (figure 29B). Amounts of binding activity between different tissues are not comparable, but the relative intensities of Oct-1 and Oct-R bands are comparable. Generally where both factors are present the intensity of Oct-1 (detected by the oct probe) and Oct-R (detected by the H2B probe) bands is approximately equal. However, with oviduct extract the Oct-1 band is much more intense than the Oct-R band.

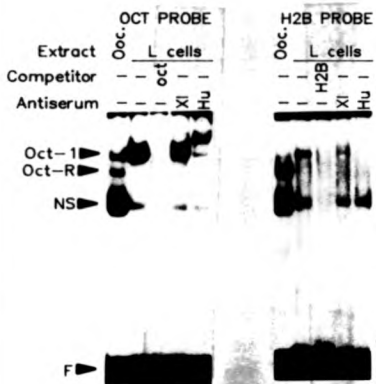
*INSERT:

The amount of these factors varies (upto 2-3 times) between stages of development, but the variation is erratic and probably due to errors in gel loading and/or equalising extract concentration.

Figure 30.

Band shift assays to show that Oct-R cannot be detected in mouse L cell extract.

Oct and H2B probes (see figure 29F) were used to maximise the binding of Oct-1 and Oct-R, respectively. Band shift assays with extract from mouse L cells (and with extract from Xenopus laevis oocytes, for comparison) are shown. X1 indicates antiserum raised against Xenopus laevis Oct-1 and Hu antiserum raised against human Oct-1 included in the binding reaction. An excess of unlabelled probe was included in the binding reaction as competitor where indicated. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



Results and Discussion

14.3 Oct-R cannot be detected in mouse L cells.

A single specific band is detected in band shift assays using octamer-containing probes and extract from mouse L cells (figure 30). Band shifts with Xenopus laevis oocyte extract are shown for comparison, and oct and H2B probes were used to maximise the binding of Xenopus laevis Oct-1 and Oct-R, respectively. The single specific band seen with L cell extract is due to Oct-1 binding as the shift is of a similar size to that produced by Xenopus laevis Oct-1 binding and since binding is inhibited by inclusion of antiserum raised against human Oct-1 (Sturm et al (1988A)) in the binding reaction and slightly inhibited by inclusion of the anti Xenopus laevis Oct-1 antiserum. Consequently a band equivalent to an Oct-R binding activity is not present in mouse L cell extract. Interestingly, mouse Oct-1 shows significant binding to the H2B probe, whereas Xenopus laevis Oct-1 does not detectably bind this probe.

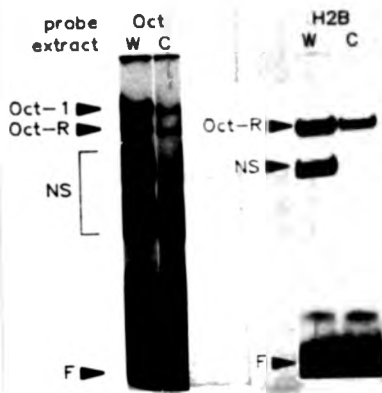
14.4 Location of Oct-1 and Oct-R in the oocyte.

Band shift assays were carried out using extract prepared from whole and manually enucleated oocytes (see figure 31). These assays (which include the equivalent of one whole oocyte or cytoplasm) indicate that approximately 25% of Oct-1 and 50% of Oct-R present in the oocyte is located in the cytoplasm (enucleated oocyte). The non-specific binding activity (which is particularly clear in the assay with the H2B probe) serves as an internal control to show that

Figure 31

Location of Oct-1 and Oct-R in Xenopus laevis oocytes determined by band shift assays.

Oct and H2B probes (see figure 28F) were used to maximise Oct-1 and Oct-R binding, respectively. W indicates whole oocyte extract and C indicates cytoplasmic (enucleated oocyte) extract. Extract from the equivalent of one oocyte or cytoplasm was used in each assay. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



Results and Discussion

enucleation without the loss of cytoplasm has been achieved. The lowest mobility non-specific binding activity (nuclear protein) is absent from the cytoplasmic extract, whereas the greatest mobility non-specific binding activity (cytoplasmic protein) is present in equal amounts in whole and cytoplasmic extract. A variable amount of Oct-1 and Oct-2 binding activity was found in isolated nuclei (not shown). This was probably due to leakage of these proteins from the nucleus during isolation of the nuclei, or because nuclei could not be obtained free of Barth's-X (the buffer in which oocytes are maintained, and hence nuclei isolated). This buffer may not maintain binding proteins in an active form.

14.5 Levels of Oct-1 and Oct-2 in cells in which DNA synthesis has been inhibited and in cells which have been serum starved.

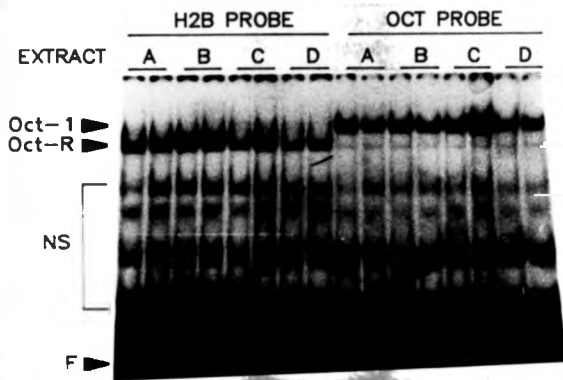
Mammalian Oct-1 has been implicated in the control of the octamer-dependent stimulation of histone H2B transcription on entry into S-phase of the cell-cycle. Consequently there has been interest in whether the binding activity of Oct-1 is cell-cycle regulated.

Xenopus laevis Xtc cells were treated with hydroxyurea, which inhibits DNA synthesis. The levels of Oct-1 and Oct-2 (determined by band shift assays, see figure 32) show no significant change following this treatment. The protein concentration of extracts from treated and untreated cells was determined and equal amounts of extract used in each

Figure 32

Band shift assays to determine the levels of Oct-1 and Oct-R in Xenopus laevis Xtc cells treated with hydroxyurea or serum-starved.

Oct and H2B probes (see figure 28F) were used to maximise Oct-1 and Oct-R binding, respectively. Duplicate flasks of cells were treated, and assays with extracts from these flasks are bracketed together. Extracts A to D were from cells treated as follows: (A) Untreated growing cells (B) Treated with 5mM hydroxyurea for 90 minutes (C) Serum-starved for 18 hours and then serum added back for 6 hours (D) Serum-starved for 18 hours. The positions of Oct-1, Oct-R and non-specific complexes (NS) and free probe (F) are indicated.



Results and Discussion

assay. Extract from duplicate flasks was assayed in each case. Our laboratory has previously shown that treatment of Xtc cells with hydroxyurea causes a rapid reduction in histone mRNA levels to 25% of that seen in untreated cells (Old et al (1985)).

Xtc cells were serum starved to block at the beginning of G1 phase of the cell-cycle, and then released from the block. The levels of Oct-1 and Oct-R (determined by band shift assays, see figure 32) show no significant change in serum starved cells relative to growing cells and adding complete medium back to starved cells resulted in no change in levels relative to starved cells. As in the case of hydroxyurea treatment, extract protein concentrations were normalised, and duplicate flasks of cells assayed.

14.6 Conclusions and speculation regarding *Xenopus laevis* Oct-1 and Oct-R based on the affinity/distribution data.

Oct-1 and Oct-R have different binding affinities. In histone H2B gene promoters the octamer motif occurs as the core of a longer consensus sequence known as the H2B box (see introduction section 1.1.2). The most common octamer motif in the *Xenopus laevis* H2B box is a 7 out of 8 match (GTTTGCAT) to the canonical octamer (ATTTGCAT) (see section 1.1.2 and figure 33). I have tested the binding affinities of Oct-1 and Oct-R for these two octamer motifs in the context of the H2B box and in the context of the mouse

Figure 33

X. laevis H2B boxes and band shift probes.

Xenopus laevis histone gene promoter H2B boxes, with core octamer motif, aligned to each other and to the corresponding region of oligos used as probes in band shift assays. The name of the Xenopus laevis histone gene cluster (Perry et al (1985), Moorman et al (1982), Aldridge (1986)) in which the H2B box is found is given in the left-hand column and the name of the band shift probe in which these sequences are found in the right-hand column. H2B, Oct and MUT probes contain the octamer in the context of the H2B box, k.H2B and k.Oct probes contain the octamer in the context of the mouse heavy chain enhancer.

OCTAMER

XIHW23	CTGCCTC	<u>GTTTGCAT</u>	GGGA	H2B PROBE
XIHW11	CTGCCTC	<u>GTTTGCAT</u>	GGGA	
XIH1	CTGCCTC	<u>GTTTGCAT</u>	GGGA	
XIHW8	<u>CAGCCTC</u>	<u>GTTTGCAT</u>	<u>GGGA</u>	
XIH3	<u>GCCAGTG</u>	<u>CTTTACAT</u>	<u>GGGC</u>	
XIHh1	CTGCCT T	<u>ATTTGCAT</u>	GGGA	
XIHW7	CTACCT T	<u>ATTTGCAT</u>	GGGA	
XIHW24	CTGCCT T	<u>ATTTGCAT</u>	GGGA	
	CTGCCTC	<u>ATTTGCAT</u>	GGGA	OCT PROBE
	CTGGGTA	<u>GTTTGCAT</u>	<u>TTC T</u>	K-H2B PROBE
	CTGGGTA	<u>ATTTGCAT</u>	<u>TTC T</u>	K-OCT PROBE
	CTGCCTC	<u>GGTTGAAT</u>	GGGA	MUT PROBE

Results and Discussion

immunoglobulin heavy chain enhancer (sequences of probes are shown in figure 33). Oct-1 binds well to the perfect octamer in either context (slightly more strongly in the H2B box) and has low affinity for the 7/8 octamer in either context. Oct-R binds both octamer motifs strongly (the 7/8 octamer more strongly) but only in the context of the H2B box. Little Oct-R binding to either motif in the context of the heavy chain enhancer is seen.

Xenopus laevis Oct-1 is similar to human Oct-1/Oct-2 in that it binds well to the perfect octamer regardless of flanking sequences. Flanking sequences are necessary to stabilise the binding of human Oct-1 to degenerate octamer motifs, but not the perfect octamer (see section 1.3.2). However, Oct-R is distinct from Oct-1 in that although binding is octamer-dependent, it does not bind well to the perfect octamer (or the 7/8 octamer) regardless of the context. Binding of Oct-R to the perfect octamer (and 7/8 octamer) is stabilised by the H2B box flanking sequences. Oct-R may be better described as a H2B box binding protein. As already mentioned the Oct-R POU domain (if DNA binding is via a POU domain) may be antigenically distinct from the human Oct-1/Oct-2 POU domain (see section 1.3.2), and this may correlate with the distinct pattern of binding. In line with this is the observation that the X. laevis Oct-1 POU domain used to probe a genomic Southern blot (chapter 7) gives a similar hybridisation pattern (at low stringency) to a probe from Oct-1 outside the POU domain. This indicates

Results and Discussion

that sequences similar to the Oct-1 POU domain do not occur in the X. laevis genome.

Mammalian Oct-1 has been implicated in the control of the octamer-dependent stimulation of histone H2B transcription on entry into S-phase of the cell-cycle. The binding properties of Xenopus laevis Oct-R suggest that it may have a role in histone H2B transcription in Xenopus laevis, although a functional assay for the activity of this factor is clearly required. Consequently I was interested to see if the distribution of Oct-R was consistent with a potential role as a ubiquitously expressed gene transcription factor.

Xenopus laevis Oct-1 is present at an approximately constant level throughout early development, and is also present in all tissues tested and in the cell line, Xtc. Oct-R has a similar distribution with the exception that it could not be detected in the protein extract from Xenopus laevis blood. The widespread distribution of both factors is consistent with their putative roles as ubiquitously expressed gene transcription factors.

A family of octamer-binding proteins (Oct-1 to Oct-10) has been identified in the mouse (see Scholer et al (1989) and introduction section 1.4.5). However, the distribution of Xenopus laevis Oct-R (which is distinct from the distribution of these mouse factors and any other previously identified octamer binding-factors) suggests that it is not a homologue of one of these factors, but a previously unidentified octamer-binding factor. In fact, a binding

Results and Discussion

activity similar to Oct-B (which has a wide distribution in Xenopus laevis) could not be detected in extract from mouse L cells. It is also interesting to note that the probe used to identify the family of mouse octamer-binding proteins, which includes proteins present in unfertilised oocytes and proteins differentially expressed at stages of mouse embryogenesis, was from the immunoglobulin heavy chain enhancer. I found that a similar probe did not detect homologues of these mouse octamer-binding proteins in Xenopus laevis oocytes and embryogenesis. However, the inability to detect homologues of these mouse octamer binding proteins should not be regarded as proof that such proteins do not exist. For example, various groups did not detect the factor Oct-2B in mammalian B cell extract (see introduction, section 1.1.1). However, as mentioned earlier, it appears that sequences closely related to the Oct-1 POU domain do not occur in the X. laevis genome.

Like Xenopus laevis Oct-1, mouse Oct-1 binds the H2B probe (7/8 octamer in H2B box) with lower affinity than the oct probe (perfect octamer in H2B box). However, mouse Oct-1 binds detectably to the H2B probe, whereas Xenopus laevis Oct-1 binding can only be demonstrated by inclusion of a large excess of the cold oligo as competitor in a binding reaction. Consequently, mouse and Xenopus laevis Oct-1 have different binding affinities, presumably due to differences within the POU domain. The sequence of mouse Oct-1 has not been reported, but since mouse and human Oct-2 are identical

Results and Discussion

within the POU domain (Hatzopoulos et al (1990)), and chicken and human Oct-1 are identical within the POU specific and POU homeo boxes (Petryniak et al (1990)) it is perhaps safe to assume that mouse and human oct-1 are identical within the POU specific and POU homeo boxes. If so, the two differences seen between human Oct-1 and Xl-Oct-1A/Xl-Oct-1B (see figure 17, the differences to the two clones are not the same) within the POU specific and POU homeo boxes (the linker between the 2 boxes is thought not to influence binding) probably account for the difference in binding affinity.

Oct-1 and Oct-2 are present at a constant level throughout early Xenopus laevis development (from oocyte to the latest embryonic stage tested, stage 26). Consequently these proteins are stored in the oocyte, the amount in an oocyte being sufficient for at least the approximately 8×10^4 cells of a stage 26 embryo (David (1965)). Early amphibian development is extremely rapid (the approximately 30 000 cells of the blastula stage are produced within 10 hours after fertilisation). Consequently, certain proteins are required in large amount. These include histones, which are required to associate with newly synthesised DNA. There are 3 well studied strategies which are used to provide sufficient histone to keep up with the high rate of DNA synthesis (reviewed in Old and Woodland (1984), Woodland (1982), Woodland et al (1983)). Firstly the Xenopus oocyte nucleus contains sufficient stored histone protein to

Results and Discussion

assemble about 20 000 nuclei. This protein is synthesised during the prolonged period of oogenesis. Secondly, histone mRNA is stored in the oocyte. After oocyte maturation translation of stored histone message begins, and provides sufficient histone for about one third of the histone present at the blastula stage. Histone mRNA in somatic cells is unstable, being rapidly degraded at the end of S-phase of the cell-cycle. In Xenopus (unlike, for example, sea urchins) stage specific histone genes are not present, and so the normally unstable message must be somehow stabilised in the oocyte. This could be by association with other factors, or alternatively by polyadenylation of histone mRNA (which is not polyadenylated in somatic cells). The bulk of histone mRNA in oocytes is polyadenylated, but becomes depolyadenylated by the unfertilised egg stage. Thirdly histone genes are reiterated to increase their overall amount of transcription (Xenopus has 50-100 copies of histone genes). Histone gene transcription first occurs at the mid blastula stage, and by the gastrula stage all histone mRNA translated is newly synthesised. In view of this it seems possible that histone gene transcription factors are stored as part of the mechanism to provide sufficient histone to keep up with the high rate of DNA synthesis observed in early development. The amount of Oct-1/Oct-R stored in the oocyte is equivalent to the the amount of RNA polymerase (I, II and III) stored, which is equivalent to the amount present in the approximately 80 000

Results and Discussion

cells of the tailbud embryo (Roeder (1974)).

Part of the Oct-1 and Oct-R store of the oocyte is located in the cytoplasm (approximately 50% of Oct-R and 25% of Oct-1). This could be to avoid effects such as squelching (reviewed in Ptashne (1988)). Activating regions of transcription factors may interact (perhaps more weakly) with target proteins without association with DNA. Consequently a high concentration of an activator could depress transcription of genes lacking the activator target site by interaction with and sequestration of general transcription factors (for example components of the RNA polymerase II transcription complex). In yeast over-production of the transcription factor Gal4 inhibits the transcription of genes lacking the Gal4 binding site (Gill and Ptashne (1988)). Squelching does not require the Gal4 DNA binding domain. Expression of a hybrid factor containing the Gal4 DNA binding domain and the VP16 (HSV transactivator, see section 1.1.5) acidic activation domain depresses genes lacking the binding site in mammalian cells, and activates genes containing the binding site (Sadowski et al (1983)). VP16 has been shown to interact with TFIID (TATA box binding factor, and component of the RNA polymerase II transcription initiation complex) (Stringer et al (1990)). The acidic activation domain of VP16 and TFIID interact in the absence of DNA, and this interaction may be the cause of the squelching observed on over-expression of VP16. Too much of the large Oct-1/Oct-R store

Results and Discussion

of the oocyte in the nucleus may interact with and hence sequester other transcription factors, leading to the depression of genes not under the control of Oct-1.

There are other examples of transcription factors which are located in the cytoplasm of some cell types. The immunoglobulin kappa chain transcription factor, NF-kB, is present in the cytoplasm of unstimulated lymphocytes (which do not express immunoglobulin genes) (Baeuerle and Baltimore (1988)). However, in B cells (which express immunoglobulin genes) the transcriptional activator is located in the nucleus. Location in the cytoplasm seems to be determined by association with a cytoplasmic anchor (I-kB). The association is probably terminated to allow nuclear localisation and transcriptional activation by protein kinase C dependent modification of the anchor protein. The dorsal gene of Drosophila, which probably encodes a transcription factor (Rushlow et al (1989), Steward (1989), Roth et al (1989), Kamens et al (1990)), is a member of a class of genes known as the dorsal group. Mutation of these genes causes dorsalisation of the embryo. The dorsal group genes act on each other, and dorsal being a probable transcription factor, is the last gene in the pathway. The amount of dorsal protein is equal along the dorso-ventral axis of the embryo, but there is a gradient of nuclear localisation. More dorsal protein is present in the nucleus of more ventral tissues. Mutation of only one gene (cactus) known to effect the dorso-ventral axis causes the embryo to

Results and Discussion

become ventralised. Cactus could be analogous to I-kB, that is a cytoplasmic anchor for dorsal, its loss causing dorsal to become nuclear and specify ventral tissues. It used to be thought that steroid hormones bound to their receptor in the cytoplasm, causing it to become translocated to the nucleus where it activated transcription. This seems to hold true for the glucocorticoid receptor, but most other receptors (eg. oestrogen, progesterone and thyroid hormone receptors) are present in the nucleus in the absence of hormone (Picard and Yamamoto (1987), Guiochon-Mantel et al (1989), and references therein). Oct-1/Oct-R could be held in the cytoplasm of the oocyte by association with an hypothetical anchor which over-rides the nuclear localisation signal. A nuclear localisation signal seems to consist of a short run of lysine/arginine residues (basic) (reviewed in Hunt (1989)), but this signal alone may not sufficient for nuclear localisation. Alternatively nuclear localisation could require modification of the protein. Oct-1 is known to be O-glycosylated and phosphorylated (Tanaka and Herr (1990), Murphy et al (1989)).

The octamer motif is required for the S-phase specific stimulation of histone H2B gene transcription (LaBella et al (1988)) and extracts from HeLa cells synchronised in S-phase of the cell-cycle are able to activate transcription in vitro, but extracts from HeLa cells synchronised in G2 are not. An Oct-1 - containing fraction from HeLa cell extract purified by affinity for the octamer DNA sequence is able to

Results and Discussion

activate histone H2B transcription in vitro (Fletcher et al (1987)). Consequently there has been interest in whether the binding activity of Oct-1 is cell-cycle regulated and I was also interested to see if the binding activity of Oct-2 is cell-cycle regulated.

There is no evidence that human Oct-1 binding activity is cell-cycle regulated, and the evidence from other species is contradictory. Synchronisation of avian cells by inhibition of DNA synthesis with aphidicolin indicated that the level of Oct-1 binding activity did not vary through the cell-cycle, although the level of another factor which binds to a promoter element responsible for the S-phase specific stimulation of histone H1 transcription was increased at S-phase (Dalton and Wells (1988A)). However, synchronisation of hamster cells by serum starvation indicated that the level of Oct-1 binding activity was increased at S-phase of the cell-cycle (Ito et al (1989)). I have found that inhibition of DNA synthesis in, or serum starvation of, Xenopus laevis cells has little effect on the levels of Oct-1 and Oct-2 binding activity relative to that seen in growing cells. It appears that the stimulation of H2B transcription on entry into S-phase of the cell-cycle may not be a result of periodic octamer factor binding, particularly since Oct-1 is also involved in the transcription of ubiquitously expressed genes which are not cell-cycle regulated (eg. snRNA). As discussed in the introduction (see section 1.2.4) the explanation to this

Results and Discussion

problem may be that there is an as yet undiscovered cell-cycle regulated modification of Oct-1 (and possibly Oct-R). Oct-1 is known to be phosphorylated and O-glycosylated (Tanaka and Herr (1990), Murphy et al (1989)). Alternatively there may be an interaction with another transcriptional regulator which is cell-cycle regulated. Oct-1 has been shown to interact with the herpes simplex virus transactivator, VP16, to activate transcription from octamer-TATA box promoters (see section 1.2.2) and transfection of VP16 stimulates the transcription of cellular H2B genes (Latchman et al (1989)). There may be a cellular analogue of VP16, which recognises the flanking sequences of the octamer in the H2B-box, interacting with Oct-1 (and possibly Oct-R) to activate histone H2B gene transcription.

The discovery of Oct-R in Xenopus laevis complicates the situation. I have not shown that Oct-R is an H2B transcription factor, or even a transcription factor at all. But, it seems likely that a protein with affinity for a functional promoter element would be a transcriptional regulator, and the facts that Oct-R has preferential affinity for the octamer motif in the H2B box, and has a widespread distribution are suggestive. If, in fact, Oct-R is an H2B transcription factor then important questions arise. Firstly, why does Xenopus have Oct-R and mammals do not? I have only analysed L cells, but Oct-R has not been reported in other species. However, this could be a

Results and Discussion

consequence of not using a probe with the correct flanking sequences. I have compared X. laevis, chicken, human and mouse H2B boxes (see introduction section 1.1.2). The core octamer motif is strongly conserved between species, however the sequences flanking the octamer motif are more strongly conserved between genes within a particular species, than between species. Oct-R, although binding is octamer-dependent, is perhaps better described as an H2B box binding protein. The equivalent factor in another species may require a probe containing the H2B box of that species to be detected. Such a probe has not been used to detect mammalian binding proteins. Secondly, why are two H2B transcription factors required in Xenopus? The evidence that human Oct-1 is a H2B transcription factor is very compelling, and since Xenopus and human Oct-1 are so well conserved it is extremely likely that Oct-1 is a H2B transcription factor in Xenopus. An explanation for the potential existence of two H2B transcription factors is further clouded by the fact that the two factors have a similar distribution, since this means that there is no apparent potential for tissue/stage specific expression of particular histone genes by virtue of the different affinity of their octamer motifs for Oct-1 and Oct-R. If the in vitro binding studies reflect the in vivo situation then Oct-1 is unable to bind strongly to a subset of H2B gene promoters (containing the 7/8 octamer), whereas Oct-R would preferentially bind these promoters. There is no obvious explanation for this pattern of binding to, and

Results and Discussion

possibly activation of particular H2B genes by Oct-1 and Oct-R.

Results and Discussion

Chapter 15.

Do the octamer motif and Oct-1/Oct-2 regulate *Xenopus laevis* histone H2B genes?

Introduction.

It was decided to try and demonstrate a role for the octamer motif in the regulation of a *Xenopus laevis* histone H2B gene, and possibly gain evidence that Oct-1/Oct-2 are H2B transcription factors.

A cloned *Xenopus laevis* histone gene cluster (gift of R. W. Old) was manipulated so that expression of H2A and H2B genes (since these genes occur as a divergently expressed pair, and in chicken the octamer motif regulates the expression of both genes. See section 1.1.2) could be distinguished from the expression of endogenous genes when introduced into oocytes by microinjection, or into cell lines by transfection. However, expression of the genes could not be detected in *Xenopus laevis* cells transfected using the standard methods used to transfect mammalian cells. I suspect that this was due to inefficient transfection, rather than that the histone gene cluster was not expressed when inside the cell. A role for the octamer motif in the regulation of *Xenopus* H2B genes has not been demonstrated, and was simply postulated by analogy to other systems. A previous study on the promoter sequences required for the expression of *Xenopus laevis* histone genes in microinjected oocytes made a deletion series through the the H2B promoter (Heindl et al (1988)), and since the CCAAT box was removed prior to the octamer, and this

Results and Discussion

caused basal levels of expression to be observed, the presence or absence of a role for the octamer motif could not be demonstrated.

15.1 Preparation of a H2A-H2B expression construct.

I was provided with a X. laevis histone gene cluster (Xlhw23), which consisted of a truncated H2B gene, and complete H2A, H3 and H4 genes, in the vector pBR325. This clone, pRW23, was a gift of R. W. Old. A restriction map and the nucleotide sequences of the coding regions, and of the H2A-H2B intergenic region were available (R. W. Old, pers comm., and Aldridge (1986)). The H2A and H2B genes occur as a divergently expressed gene pair, and the octamer motif is the 7/8 octamer (GTTTGCAT) which is the most common octamer seen in Xenopus laevis H2B boxes (see figure 33). This octamer binds well to Oct-R in vitro, but does not detectably bind Oct-1 (see section 14.1). The H2A-H2B intergenic region is shown in figure 34A, and a map of Xlhw23 in figure 34B.

The first step in making an expression construct (illustrated in figure 34B) was to insert an internal control reference gene into pRW23. This is the 'albino' gene, which consists of the 5' end of a Xenopus laevis albumin gene (670 bases of the 5' flanking sequence and 50 bases downstream of the transcription initiation site) fused to the 3' end of a Xenopus laevis histone H3 gene (370 bases consisting of a little of the 3' end of the gene and some 3' flanking sequence). This fusion gene was made in our lab, and is

Figure 34.

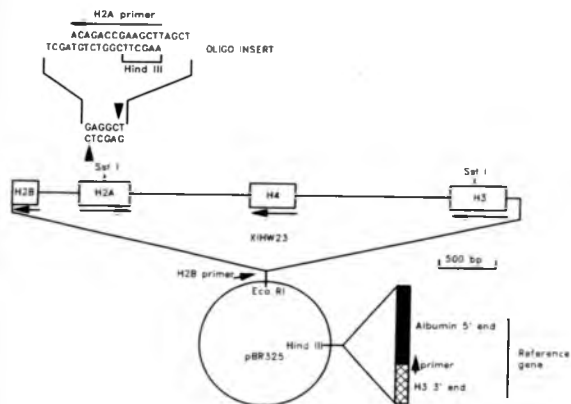
X1HW23 H2A-H2B intergenic region and structure of
pH2A/B.exp.

- (A) X1HW23 H2A-H2B intergenic region (Aldridge (1986)).
Position of the H2B box with core octamer is shown.
Position of predicted CAP, TATA and CCAAT sites is also
shown.
- (B) Diagram to illustrate the modification of pRW23 to
create pH2A/B.exp. The 'albino' reference gene was
inserted into the Hind III site of pBR325, and the
duplex oligonucleotide shown into the Sat I site of the
H2A gene. The location of primers used for primer
extension analysis of transcripts from the construct are
indicated. The H2A primer is an 18mer, and albino and
H2B primers are 17mers.

A

GGA TTTCAACAAGACTCAGTATCTTTTGGCTATACACTGTACCCAAACTGTGGGGTCACTCTTTTATAGTC
 H2B START CAP TATAA
 CTCCCATGCCAAAGAGGCGAGCTCTGTCTTGAGTTTTCATTTGCCAATTCTGTCACTAGAGCTAAAGCG
 H2B BOX OCAAT
 OCTAMER
 CTACTGCCAAGACTAAGTCAAGGAGAACATAGGGATATTATCCAATCAAGCCAAAGATTAGGCTGT
 GATATTTTGGTGTCTTCTGTTAATACAGCTCTGGCTATGGCTGCTTGAAGCTACCCCAATCAGCGAAC
 CCAAT
 GGTCAGCTCTATAAATGAGCGATCCGGGGCTTGGTGGTTATGCTAGATTGCTACTAGAAATTCGACAAC ATG
 TATAA CAP H2A START

B



Results and Discussion

described in Old et al (1988). The source of the gene was a Hind III fragment from albino Δ -670 (gift of A. R. Brooks) (Old et al (1988)). The Hind III fragment was inserted into the unique Hind III site of pRW23. The gene has been shown to express on injection into oocytes, and transcripts can be detected by primer extension using the albino primer, which extends over the junction of the 2 components of the fusion gene.

It was intended to assay H2A and H2B transcription by primer extension analysis of RNA from injected oocytes. The H2B gene is truncated, and consequently a primer homologous to the pBR325 sequence (H2B primer, see figure 34B) can be used to distinguish transcripts of introduced genes from those of endogenous genes. However, to distinguish transcripts of the introduced H2A gene from those of endogenous genes, it was necessary to insert a duplex oligonucleotide into the H2A gene of pRW23. The oligonucleotide (illustrated in figure 34B) consists of Sat I sticky ends (but on insertion into a Sat I site, sites are not reformed, as the base adjacent to the sticky end is not that found in an Sat I site) and an off-centre Hind III site. Two Sat I sites occur in (pRW23+albino), one in the H2A gene and one in the H3 gene. A (pRW23+albino) partial Sat I digest was carried out, and a fragment corresponding to the linear clone (ie. only one Sat I site cut) was isolated. The duplex oligonucleotide was ligated into this fragment. Recombinants were identified by the loss of one Sat I site, and the position of the oligo

Results and Discussion

insert was determined by the size of a fragment released by Hind III digestion (the oligo contains a Hind III site). Hind III digestion releases the albino gene and a fragment from the oligo insert to the Hind III site of pBR325. If the oligo insert is in the H3 gene an approximately 1.7kb fragment is released, whereas if the insert is in the H2A gene a fragment of approximately 4.6kb is released. Orientation of the oligo insert in the H2A gene was determined by labelling the products of a Hind III - Sau 3A double digest by end filling, and visualising them by autoradiography after resolution on a denaturing polyacrylamide gel. A Sau 3A site occurs approximately 50 bases from the Sat I site in the H2A gene. Since the position of the Hind III site in the oligo is off-centre, a Hind III - Sau 3A digest will yield a fragment of 52 or 60 bases, depending on the orientation of the oligo insert. One strand of the duplex oligonucleotide (depending on the orientation of the oligo insert) was used as a primer for primer extension analysis (H2A primer). This construct was called pH2A/B.exp. The predicted sizes of primer extension products (using the predicted CAP sites shown in figure 34A) are approximately 240 bases for the H2B gene, and 195 bases for the H2A gene. The 'albino' gene gives a product of around 50 bases. Consequently it should be possible to analyse transcription of all 3 genes in a single reaction.

Results and Discussion

15.2 Expression of pH2A/B.exp can be detected in microinjected oocytes.

pH2A/B.exp was microinjected into the nuclei of Xenopus laevis oocytes, and the injected oocytes maintained overnight. RNA was made from surviving oocytes and transcript levels from introduced 'albone', H2A and H2B genes was assayed by primer extension analysis. The result is shown in figure 35. The reference gene gives the previously reported product, and H2A and H2B genes products are in the predicted size range. It is not apparent here, but the H2B often gives two transcripts (around 239 and 243 bases) indicating that two transcription initiation sites are used. As a control, the primers detect no transcripts in un-injected oocytes. Mixture of all 3 primers shows that expression of all 3 genes can be monitored in a single reaction. Expression of this clone, as already mentioned, could not be detected in Xenopus laevis Xtc cells transfected using standard methods used for mammalian cells (data not shown).

15.3 An attempt to modulate pH2A/B.exp expression by competition with oct factor binding sites.

The band shift probes described in chapter 18 (see figure 28F) provide a set of binding sites which (1) bind only Oct-1 strongly (k.oct probe), (2) bind only Oct-R strongly (H2B probe), (3) bind Oct-1 and Oct-R (oct probe), (4) bind neither Oct-1 or Oct-R strongly (k.H2B and mut probes). Consequently, these oligos were co-injected in molar excess

Figure 35.

pH2A/B.exp is expressed in micro-injected oocytes.

Approximately 10nl of pH2A/B.exp (0.5 mg/ml) was injected into the nucleus of X. laevis oocytes. Oocytes were maintained overnight, and RNA prepared from surviving oocytes. RNA from the equivalent of 4 oocytes was hybridised with 200pg of end-labelled oligonucleotide primer. Hybridisation to RNA from un-injected oocytes is included as a control. After hybridisation, primer extension with reverse transcriptase was carried out, and the products resolved on an 8% denaturing polyacrylamide gel. An autoradiograph of the gel is shown. H2A, H2B and albino primers (see Figure 34) were used where indicated. The primer extension products from transcripts of the H2A, H2B and 'albino' genes are indicated. Size markers are pBR322 digested with Hpa II and labelled by end-filling.

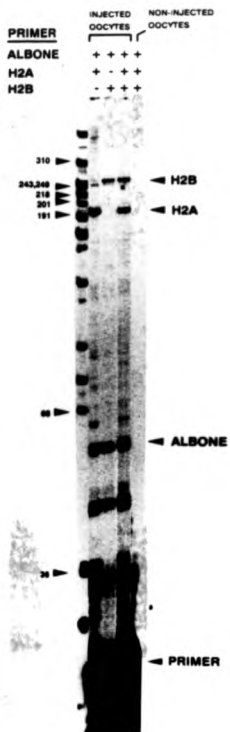


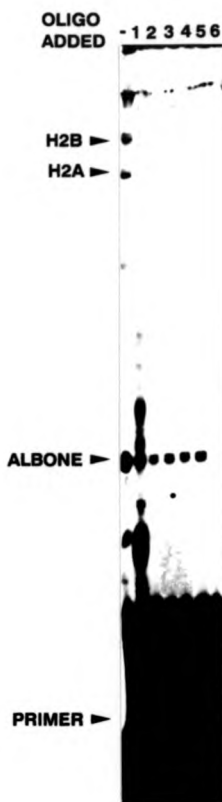
Figure 36.

Effect of competition with Oct factor binding sites on pH2A/B.exp expression in oocytes.

Approximately 10nl of pH2A/B.exp (1mg/ml) was mixed with an equal volume of concatemerised oligonucleotide (1mg/ml) or water (~ track) and injected into the nuclei of oocytes. Primer extension analysis was performed as described in figure 35. A mixture of H2A, H2B and albino primers was used. Position of primer extension products from transcripts of H2A, H2B and 'albino' genes are indicated.

Oligonucleotide sequences are given in figure 28F, and the numbers refer to mixing of pH2A/B.exp with the following oligonucleotides:

1. non-specific oligonucleotide used in band shift assays.
2. Oct oligo band shift probe.
3. H2B oligo band shift probe.
4. K.oct oligo band shift probe.
5. K.H2B oligo band shift probe.
6. Mut oligo band shift probe.



Results and Discussion

with pH2A/B.exp in an attempt to modulate H2A/H2B expression by competing for Oct-1/Oct-R binding. If competition were seen it would indicate the relative contribution of Oct-1/Oct-R to transcriptional activation of the H2A/B genes contained in this construct. The oligonucleotides were concatemerised to increase their stability in oocytes. The non-specific duplex oligonucleotide used in band shift assays was included as a control competitor. After injection transcript levels were monitored by primer extension. The result is shown in figure 36. The result of this experiment cannot be interpreted since all the oligonucleotides caused a large non-specific reduction in transcription of all 3 genes, including the reference gene.

15.4 Making a mutation in the octamer motif associated with the H2B gene of pH2A/B.exp.

Following the lack of success with oligonucleotide competition it was decided to try and demonstrate a role for the octamer motif in the control of H2A/H2B expression by introducing a mutation into the octamer motif. The mutation made (illustrated in figure 37A) converts the 7/8 octamer motif of X1hw23 to the sequence of the mut band shift probe (see chapter 14) which shows much reduced Oct-R binding activity.

The Eco RI fragment, containing the histone gene cluster, from pH2A/B.exp was subcloned into M13mp18, and oligonucleotide mutagenesis used to convert the octamer in

Figure 37.

Mutagenesis of the octamer motif associated with the H2B gene of pH2A/B.exp.

- (A) The Eco RI insert from pH2A/B.exp was subcloned into M13mp18, and the oligonucleotide indicated used to introduce a double point mutation into the H2B box of the H2B gene.
- (B) Screening of potential H2B octamer mutants. 1ul and 5ul of M13 phage stock from 6 potential mutants and from the original non-mutant subclone was spotted onto duplicate nitrocellulose filters, denatured, baked and the nucleic acids on the filters hybridised to end-labelled mutagenic oligonucleotide in 6x SSC at room temperature. One filter was washed at 37°C (Tm-17) and one filter at 54°C (Tm-2) in 6x SSC.

A

MUTAGENIC OLIGONUCLEOTIDE

H2B BOX

GCCTC	GTTGAAT	GGGA
	* *	
CTGCCTC	GTTGCAI	GGGA
	OCTANNE	

B

WASHING
TEMPERATURE
(°C)

37(Tm-17)

52(Tm-2)

VOLUME
PHAGE STOCK
(MICROLITRES)

1 5

1 5

POSSIBLE
MUTANTS

1

2

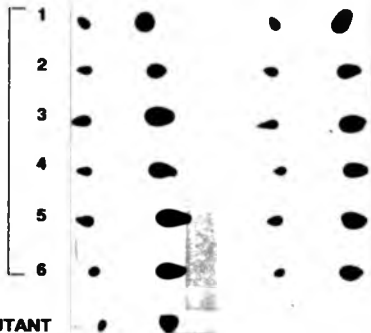
3

4

5

6

NON-MUTANT



Results and Discussion

the H2B promoter (GTTTGCAT) to GGTGAAT. Potential mutants were screened by hybridisation of single stranded phage stock to end labelled mutagenic oligonucleotide. The melting point (T_m) of the perfectly matched mutant duplex was calculated to be 54°C. Six potential mutants were screened, along with the original non-mutant clone as control. The result is shown in figure 37B. All clones hybridised to probe after washing at 37°C. All putative mutants hybridised to probe after washing at T_m-2 (52°C), but the non-mutant control did not hybridise at this temperature. Consequently the frequency of mutants obtained was 100%. One mutant was selected, and checked by sequencing. The Eco RI fragment from the mutant clone was inserted back into the parent vector to create pH2A/B.exp-mut.

15.5 Effect of the octamer mutation on H2A/B expression.

The wild-type and mutant octamer-containing constructs were microinjected into the nuclei of oocytes, and levels of H2A, H2B and albino gene expression determined by primer extension analysis of RNA. The result is shown in figure 38A. Laser scanning densitometry of the autoradiograph indicates that (relative to the 'albino' reference gene, expression of which is a roughly twice as great in oocytes injected with H2A/B.exp-mut) H2A expression is reduced to 40% of that seen with the wild-type octamer, and H2B expression to 60% by mutation of the octamer motif. This effect is small, but reproducible. A repeat of this experiment is shown in figure

Figure 38.

A comparison of pH2A/B.exp and pH2A/B.exp-mut expression in oocytes.

- (A) Approximately 10nl of pH2A/B.exp (0.5mg/ml) or pH2A/B.exp-mut (0.5mg/ml) was injected into the nuclei of oocytes. Primer extension analysis was performed as described in figure 35. A mixture of H2A, H2B and albome primers was used. Position of primer extension products from H2A, H2B and albome primers are indicated.
- (B) A repeat of the experiment described in (A).

A CONSTRUCT
INJECTED

pH2A/B. asp
pH2A/B. asp-MUT



▲ H2B
▲ H2A

▲ ALBONE

B CONSTRUCT
INJECTED

pH2A/B. asp
pH2A/B. asp-MUT



▲ H2B
▲ H2A

▲ ALBONE

Results and Discussion

38B. The result is roughly the same, except that H2B expression was lowered to 40% and H2A expression to 50% by the mutation. The effect is not in line with the 5 fold difference in expression reported with wild-type and mutant human promoter constructs transfected into somatic cells (see section 1.1.2). However, if significant this is the first reported demonstration that the octamer motif is required for the transcription of Xenopus H2B genes. Also, the result is in line with that observed for chicken H2A-H2B divergent gene pairs (see section 1.1.2), since the octamer motif seems to be required for the expression of both H2B and H2A genes.

The result is also suggestive of a role for Oct-R in the transcriptional regulation of H2A/H2B genes, since in vitro Oct-1 from oocyte extract does not detectably bind this promoter, whereas Oct-R binds strongly, and the mutation made causes a large reduction in Oct-R binding.

The fact that only a small reduction in expression, as a result of the mutation, was observed could be a consequence of several factors. For example, the mut oligo (which has the same sequence as the mutant described here) shows low, but readily detectable, Oct-R binding activity in band shift assays (see chapter 14) and this may be significant. Also, a problem which arises with all oocyte injection experiments is that there may be insufficient binding factor to saturate injected genes, and so the majority of expression detected is basal level expression which would not be affected by promoter mutations.

Results and Discussion

This experiment is not definitive, and should be extended further. For example a mutation in the octamer which has a greater effect on Oct-R binding would be useful, as would a mutation which converts the octamer to an Oct-1 binding site.

It should be borne in mind that the oocyte is a special case of histone gene regulation. In Xenopus, (unlike, for example, sea urchin where developmental stage specific histone genes occur), the same histone genes are expressed in the oocyte independently of DNA synthesis, and in somatic cells tightly coupled to DNA synthesis (Old et al (1985), Woodland et al (1984), Perry et al (1986)). Unfortunately the effect of the octamer motif mutation could not be tested in somatic cells following transfection. However, by analogy to mammalian H2B genes it seems likely that the octamer motif is also required for H2A/H2B expression in somatic cells. Oct-1 and Oct-R binding activities do not vary through the cell cycle, and consequently there must be a cell-cycle regulated modification/interaction with another factor in somatic cells (see sections 1.2.4 and 14.6). Perhaps this modification/interaction is constitutive in oocytes to allow replication independent expression of H2A/H2B genes to be stimulated by octamer binding factors.

Results and Discussion

Chapter 16.

The anti Oct-1 polyclonal antiserum detects proteins other than Oct-1 on Western blots.

16.1 The anti Oct-1 antiserum specifically detects two proteins in ovary protein extract.

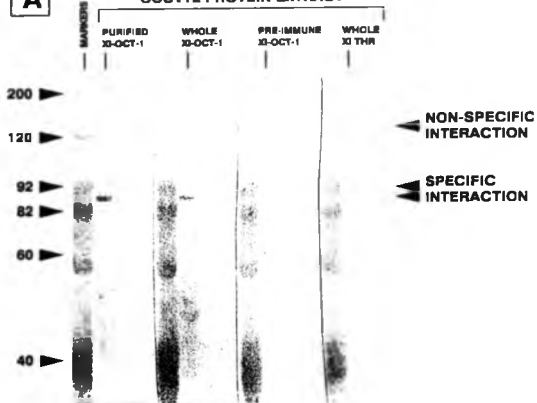
The anti Oct-1 antiserum, the preparation of which is described in chapter 11, was used to stain Western blots of ovary protein extract. The antiserum was raised against a fusion protein, produced in E. coli, which spans amino acids 16 to 231 at the N terminal end of X1-Oct-1A. As described in chapters 13 and 14 the antiserum reacts with both X. laevis and mouse Oct-1 in band shift assays. For use in staining Western blots, the antiserum was purified by affinity for the fusion protein electrophoretically transferred to a nitrocellulose filter. Antibodies bound to Western blots were detected by binding a biotinylated anti rabbit immunoglobulin, then binding streptavidin alkaline phosphatase conjugate, and then colour development with the substrates BCIP and NBT.

Figure 39A shows Western blots of oocyte protein extract stained with whole anti Oct-1 antiserum and purified anti Oct-1 antiserum, and as controls anti Oct-1 pre-immune serum and anti X. laevis thyroid hormone receptor (Gift of R. W. Old). The anti Oct-1 antisera (whole and purified) detect three protein bands: 130kD, 95kD and 85kD (the 95kD band is faint, but its presence is clearer on subsequent blots). The control antisera detect the highest molecular

Figure 39.

Anti Oct-1 antiserum specifically detects two proteins in oocyte extract.

- (A) Extract equivalent to 1 oocyte was electrophoresed in 4 tracks of a 10% separating SDS polyacrylamide gel. The proteins were electrophoretically transferred to a nitrocellulose filter. The filter was cut into strips, and each of the 4 strips incubated with the antiserum indicated (1 in 100 dilution, in TBS). Bound antibodies were detected with an Amersham blotting detection kit (biotinylated anti rabbit immunoglobulin, followed by streptavidin alkaline phosphatase conjugate, followed by colour development with the substrates BCIP and NBT). The position and molecular weight of protein markers is shown. Pre-stained molecular weight markers were used, and as these do not run true to their molecular weight they were used as positional markers (the molecular weight to which the bands correspond being determined by running adjacent to regular molecular weight markers on a Coomassie stained gel (not shown)). Two protein bands are detected as a result of specific interaction of with the anti Oct-1 antiserum (these bands are not detected by pre-immune serum and anti thyroid hormone (THR) antiserum), and also one protein band as a result of non-specific interaction (detected by all 4 antisera). The position of these bands is indicated.
- (B) Extract equivalent to 1 oocyte was electrophoresed and electroblotted as described above. The filter was stained with a 1 in 100 dilution of anti human Oct-1 antiserum, and bound antibodies detected as described above. The human Oct-1 antiserum detects the same non-specific and specific bands as the *X. laevis* anti Oct-1 antiserum. The location of these bands is indicated.

A**ANTISERUM USED TO STAIN
OOCYTE PROTEIN EXTRACT****B****OOCYTE PROTEIN STAINED
WITH ANTI HUMAN OCT-1**

Western blot showing a single lane of oocyte protein stained with anti-human OCT-1. The lane shows a broad band at the top labeled 'NON-SPECIFIC INTERACTION' and a distinct band at the bottom labeled 'SPECIFIC INTERACTION'.

Results and Discussion

weight band (130kD), and so this band is disregarded as being the result of a non-specific interaction. The smaller bands are not detected by the control antisera, and are as a result of specific interaction of proteins with the anti Oct-1 antiserum. Detection of protein bands by whole and affinity purified anti Oct-1 antiserum was similar, however, purified antiserum was routinely used for staining Western blots.

Figure 39B shows a Western blot of X. laevis oocyte protein extract, stained with anti human Oct-1 polyclonal antiserum (Gift of W. Herr, see Sturm et al (1988A)). This antiserum detects the two specific protein bands, and is consequently further evidence that a specific interaction is being detected.

The size of the two specific protein bands detected is 85kD and 95kD. This is larger than the predicted size of the Xl-Oct-1A product (79kD) and the size of the product of Xl-Oct-1A synthetic message expressed in oocytes (which is the predicted size, see chapter 9). Consequently, it seems possible that the antiserum is not detecting Oct-1, but antigenically related proteins, although the difference from the predicted size could be due to some modification.

Results and Discussion

16.2 Oct-1 and the proteins to which the anti Oct-1 antiserum reacts on a Western blot can be separated on a sucrose gradient.

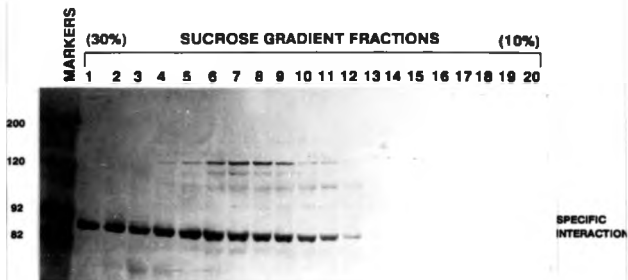
Protein extract from the equivalent of 100 oocytes was loaded onto an 11ml 30% (w/v) to 10% sucrose gradient, and the gradient centrifuged at 37000rpm and 4°C for 45 hours. The gradient was fractionated into 20x 0.5ml fractions. Every other fraction was assayed for Oct-1 (and Oct-2) binding activity by band shift assays, and proteins in each fraction were resolved by SDS-PAGE, electroblotted to nitrocellulose and stained with anti Oct-1 antiserum. The results are shown in figure 40.

The Western blot was from a shorter gel run, and the two specific bands are not resolved. The specific protein bands are very faintly detected by the anti Oct-1 antiserum in fraction 13, and then detected in each fraction to the bottom of the gradient (fraction 1). The proteins peak in fractions 5 to 7. The band shift assay indicates that Oct-1 (and Oct-2) binding activity peaks in fraction 13. Consequently, the gradient separates Oct-1 binding activity and the proteins which are specifically detected by the anti Oct-1 antiserum on a Western blot. The antiserum does react to Oct-1 in band shift assays, but this result tends to suggest that the antibody is reacting to proteins other than Oct-1 on a Western blot. The size of the proteins detected also indicated this, and as will be described later the proteins detected by the antiserum and Oct-1 have distinct

Figure 40.

The proteins to which the anti Oct-1 antiserum reacts and Oct-1 can be separated on a sucrose gradient.

- (A) Protein extract from the equivalent of 100 oocytes was fractionated on an 11ml 10 to 30% (w/v) sucrose gradient by centrifugation at 4°C and 37000rpm for 45 hours. The gradient was split into 20x 0.5ml fractions. 20ul of each fraction was resolved by SDS-PAGE, electroblotted to nitrocellulose and stained with affinity purified anti Oct-1 antiserum as described in figure 39. The position of protein bands detected as a result of specific interaction of the antiserum are indicated (two specific bands are present, but not resolved on this short gel run).
- (B) Band shift assays with the oct probe (see figure 29F). 5ul of the sucrose gradient fraction indicated was used in each assay. The positions of Oct-1, Oct-R and non-specific (NS) complexes and free probe (F) are indicated. Oct-1 and Oct-R binding activity peak in fraction 13.

A**WESTERN BLOT****B****BAND SHIFT ASSAY**

Results and Discussion

distributions in the oocyte. The apparent lack of detection on a Western blot is probably result of Oct-1 being of too low abundance. The proteins detected by the Oct-1 antiserum on a Western blot, are antigenically related to the N terminal end of Oct-1. The N terminal domain of human Oct-1 is interchangeable with the N terminal domain of Oct-2, and deletion analysis of Oct-2 indicates that the N terminal Q rich region is a transcriptional activation domain (see section 1.3.4). Proteins related to the N terminal activation domain of Oct-1/Oct-2 may occur may occur in X. laevis oocytes.

16.3 Distribution of Oct-1 related proteins in oogenesis and early development.

Protein extract from stages of oogenesis and early development were resolved by SDS-PAGE, electroblotted to a nitrocellulose filter and stained with an anti Oct-1 antiserum. The result is shown in figure 41. The equivalent of one oocyte/egg/embryo was analysed in each case. Protein extract from the cell line, Xtc, was also included.

The Oct-1 related proteins are first detected in stage II oocytes (early vitellogenesis), and reach peak levels in stage III oocytes (pigment formation apparent, but no animal-vegetal division). The level of the Oct-1 related proteins is constant throughout early development, from mature oocyte to the latest stage tested, stage 32. The Oct-1 related proteins are synthesised during oogenesis, and stored

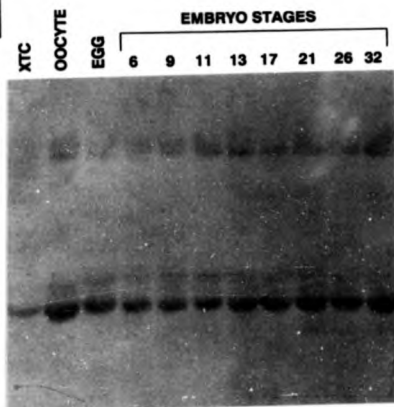
Figure 41.

Distribution of Oct-1 related proteins in oogenesis and early development.

- (A) Extract from the equivalent of one oocyte was resolved by SDS-PAGE, electroblotted to a nitrocellulose filter and stained with affinity purified anti Oct-1 antiserum as described in figure 39. The position of Oct-1 related protein bands (as a result of specific interaction with the antiserum) is indicated. Protein extract from oocytes of the stage (Dumont (1972)) indicated was analysed.
- (B) Extract from the equivalent of one mature oocyte, one egg, one embryo of the stage indicated (Nieuwkoop and Faber (1956)), and from the cell line Xtc was resolved and stained with antiserum as described above. The position of Oct-1 related protein bands (as a result of specific interaction with the antiserum) is indicated.

A

▲ SPECIFIC
▲ INTERACTION

B

▲ SPECIFIC
▲ INTERACTION

Results and Discussion

in the oocyte, the amount stored being equivalent to at least the roughly 100 000 cells present in a stage 32 embryo (Dawid (1965)). As described in chapter 14, certain proteins, which are required in large amount during early development, are synthesised during oogenesis are stored to provide sufficient protein for the extremely rapid early development phase. The proteins are also detected in the Xtc cell line.

16.4 Location of the Oct-1 related proteins in the cell.

Protein extract from the equivalent of one whole oocyte, one oocyte cytoplasm (manually enucleated oocyte) and one oocyte nucleus was resolved by SDS-PAGE, electroblotted to a nitrocellulose filter and stained with anti Oct-1 antiserum. The result is shown in figure 42A. The Oct-1 related proteins are located entirely in the oocyte cytoplasm (no protein is detected in the oocyte nuclear extract, and the amount of protein present in one cytoplasm is equivalent to the amount of protein present in one whole oocyte). Since only 25% of Oct-1 is located in the oocyte cytoplasm (see chapter 14) this is further evidence that the antiserum is reacting to proteins other than Oct-1 on Western blots.

Protein extract was prepared from whole Xtc cells, and the nuclei of Xtc cells. To confirm that the nuclear preparation contained nuclei, and no cytoplasm, nucleic acid was prepared from the nuclear and whole cell Xtc extract. The nucleic acids were electrophoresed on an ethidium-stained agarose gel. The result is shown in figure 42B. The whole cell Xtc

Figure 42.

Location of Oct-1 related proteins in the cell.

- (A) Extract from the equivalent of one whole oocyte (W), one oocyte cytoplasm (enucleated oocyte, C) and one oocyte nucleus (N) was resolved by SDS-PAGE, electroblotted to a nitrocellulose filter and stained with affinity purified anti Oct-1 antiserum as described in figure 39. The position of Oct-1 related protein bands (as a result of specific interaction with the antiserum) is indicated.
- (B) Nucleic acid prepared from whole cell Xtc extract (W) and nuclear extract (N) was electrophoresed on a 1% non-denaturing agarose gel, stained with ethidium bromide. This was to check the nuclear preparation. Chromosomal DNA (nuclear) and ribosomal RNA (cytoplasmic) bands are indicated. This demonstrates that the nuclear preparation contains nuclear components, and no cytoplasm.
- (C) Nuclear (N) and whole cell (W) Xtc cell extract was resolved and stained with antiserum as described in (A). The position of Oct-1 related protein bands (as a result of specific interaction with the antiserum) is indicated.

A

OOCYTE
W N C



SPECIFIC
INTERACTION

W-WHOLE CELL
N-NUCLEUS
C-CYTOPLASM

B

XTC CELL
N W

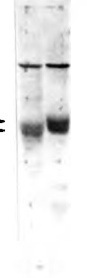


CHROMOSOMAL
DNA

RIBOSOMAL
RNA

C

XTC CELL
W N



SPECIFIC
INTERACTION

Results and Discussion

extract contains chromosomal DNA (nuclear) and ribosomal RNA (cytoplasmic), whereas the nuclear extract contains only chromosomal DNA. Nuclear and whole cell Xtc extract, from the equivalent of 200 000 cells, was resolved by SDS-PAGE, electroblotted to a nitrocellulose filter and stained with anti Oct-1 antiserum. The result is shown in figure 42C. The Oct-1 related proteins are entirely nuclear in Xtc cells.

In summary, proteins antigenically related to the N terminal (transcriptional activation) domain of Oct-1 are synthesised during oogenesis, stored in the oocyte cytoplasm (the amount equivalent to at least 100 000 somatic cells), and in somatic cells are located entirely in the nucleus.

There are other examples of proteins being stored in the oocyte cytoplasm, and subsequently becoming translocated to the nucleus of somatic cells. For example, the X. laevis c-myc protein is accumulated during oogenesis (an amount equivalent to that found in 400 000 somatic cells), and stored in the oocyte cytoplasm (Gusse et al (1989)). After fertilisation the store rapidly migrates to the nucleus.

Results and Discussion

Chapter 17.

General discussion and conclusions.

I have isolated and completely sequenced two X. laevis homologues of the human octamer-binding transcription factor, Oct-1. The degree of relatedness of the two homologues (93% similar in nucleotide and predicted amino acid sequence) indicates that these are likely to be copies of the same gene, which arose during the theoretical genome duplication event in X. laevis evolution. X. laevis and human Oct-1 display strong evolutionary conservation (85% similar in predicted amino sequence over a stretch of 750 amino acids), which presumably means that the X. laevis homologue has a similar, if not identical function to human Oct-1. Homology between human and X. laevis does, however, break down shortly before the N terminal end, at a point where alternate splicing is known to occur in human Oct-1 (W. Herr, pers. comm.). The full length X. laevis cDNA clone which I have isolated may represent a novel alternately spliced form of Oct-1.

The X. laevis Oct-1 clone which contains a complete coding sequence (Xl-Oct-1A) has been shown by expression of of synthetic transcripts in micro-injected oocytes to contain an open reading frame of the size predicted from the nucleotide sequence. This clone has also been demonstrated to encode a functional octamer-binding protein (which requires no other factors to bind DNA) by use of Oct-1 fusion proteins (expressed in E. coli) in band shift assays. The observation

Results and Discussion

that shortened an full length octamer-binding fusion proteins when mixed do not form a band corresponding to a heterodimer in band shift assays indicates that Oct-1 binds as a monomer.

Two octamer-binding proteins have been identified (in band shift assays) in X. laevis oocyte, embryo and tissue extract. A polyclonal antiserum raised against X. laevis Oct-1, and comparison to the mobility of full length Oct-1 fusion proteins in band shift assays have been used to identify one of these proteins as Oct-1. The second, previously unidentified octamer-binding protein has been termed Oct-R, for octamer-related. Oct-1 and Oct-R have different binding affinities. Oct-R binds to the consensus octamer motif (ATTGCGAT) and a degenerate octamer motif (GTTGCGAT), which is the most common octamer motif found in X. laevis histone H2B gene promoters, only in the context of the H2B promoter, and not in another sequence context. Oct-1 does not bind the degenerate octamer motif in either context, but does bind to the consensus octamer motif in either sequence context. The binding properties of Oct-R are suggestive of a role in the regulation of histone H2B transcription (although no direct evidence has been obtained), and consistent with this possibility is the fact that Oct-R has a widespread tissue distribution.

Two lines of evidence suggest that the Oct-R POU domain (if, in fact, Oct-R binds to DNA via a POU domain) may be somewhat diverged from the Oct-1 POU domain. Firstly, an antibody raised against a full-length human Oct-1 fusion protein

Results and Discussion

(which reacts to Oct-2, presumably via the conserved POU domain) does not react to Oct-R. Secondly, the X. laevis POU domain used to probe a genomic Southern blot gives a similar hybridisation pattern (at low stringency) to a probe from Oct-1 outside the POU domain. Which indicates that sequences similar to the Oct-1 POU domain do not occur in the X. laevis genome.

Since Oct-1 is believed to stimulate the S-phase specific induction of histone H2B gene transcription (see section 1.1.2), the possibility that Oct-1 binding activity is cell-cycle regulated is of interest. In mammals it is not clear if Oct-1 binding activity is cell-cycle regulated. X. laevis Oct-1 (and Oct-R) binding activity does not appear to be cell-cycle regulated.

Oct-1 and Oct-R are stored in the oocyte, in an amount equivalent to at least 80 000 somatic cells. Histone protein and message are stored in the oocyte as part of the mechanism to provide enough histones to keep-up with the high rate of DNA synthesis in early Xenopus development. It is possible that histone gene transcription factors are stored for the same purpose. Part of the Oct-1 and Oct-R store is located in the cytoplasm of the oocyte.

Oct-R cannot be detected in mouse L cells. I have compared X. laevis, chicken, human and mouse H2B boxes. The core octamer motif is strongly conserved between species, however the sequences flanking the octamer motif are more strongly conserved between genes within a particular species, than

Results and Discussion

between species. Oct-R, although binding is octamer-dependent, would perhaps be better described as an H2B box binding protein. Mammals may have an 'Oct-R' with affinity for a species-specific H2B box. Such a probe has not been used to detect mammalian binding proteins.

By mutation of the octamer motif in the promoter of a X. laevis histone H2B gene promoter I have tentatively concluded that the octamer motif is required for the expression of H2B genes (independently of DNA synthesis) in the oocyte. The H2B gene occurs, as part of a divergently expressed gene pair, with a H2A gene. The octamer motif may be required for the expression of both H2B and H2A genes. The degenerate octamer motif contained in this H2B promoter does not bind efficiently to Oct-1 in vitro, but binds well to Oct-R. Consequently, this indirectly suggests that Oct-R is required for the expression of the H2B gene.

A polyclonal antiserum raised against the N terminal domain of X. laevis Oct-1 reacts to proteins other than Oct-1 on Western blots of oocyte and embryo extract. These proteins, which are antigenically related to the N terminal domain of Oct-1, are entirely located in the cytoplasm of the oocyte, and entirely located in the nucleus of somatic cells. These proteins are synthesised during oogenesis, and stored in the oocyte in an amount equivalent to at least 100 000 somatic cells.

REFERENCES.

Aldridge, T. C. 1986. Xenopus laevis H2A and H2B genes: organisation, structure and expression. M. Sc. thesis, University of Warwick.

apRhys, C. M. J., D. M. Ciuto, E. A. O'Neill, T. J. Kelly and G. S. Hayward. 1989. Overlapping octamer and TAATGARAT motifs in the VF65-response elements in herpes simplex virus immediate-early promoters represent independent binding sites for cellular Nuclear Factor III. J. Virol. 63:2798-2812.

Artishevsky, A., S. Wooden, A. Sharma, E. Resendez Jr. and A. S. Lee. 1987. Cell-cycle regulatory sequences in a hamster histone promoter and their interaction with cellular factors. Nature 328:823-827.

Baeuerle, P. A., and D. Baltimore. 1988. I κ B: a specific inhibitor of the NF- κ B transcription factor. Science 242:540-546.

Barberis, A., G. Superti-Furga and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. Cell 50:347-359.

Barberis, A., G. Superti-Furga and M. Russlinger. 1989. The testis-specific octamer-binding protein of the sea urchin has a molecular weight of 85kDa. *Nucleic Acids Res.* 17:7114.

Baumruker, T., R. Sturm and W. Herr. 1988. OBF100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. *Genes Dev.* 2:1400-1413.

Belcourt, M. F., and P. J. Farabaugh. 1990. Ribosomal frameshifting in the yeast retrotransposon: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* 62:339-352.

Benton, W. D., and R. W. Davis. 1975. Screening λ gt recombinant clones by hybridisation to single plaques in situ. *Science* 196:180-182.

Bergman, J., B. Strich, H. Sharir, R. Ber and R. Laskov. 1990. Extinction of Ig gene expression in myeloma x fibroblast somatic cell hybrids is accompanied by repression of the oct-2 gene encoding a B-cell specific transcription factor. *EMBO J.* 9:849-855.

Biggin, M. D., and R. Tjian. 1989. Transcription factors and the control of Drosophila development. *Trends Genet.* 5:377-383.

Bodner, M., J-L. Castrillo, L. E. Theill, T. Deerinck, M. Ellisman, and M. Karin. 1988. The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. *Cell* 55:505-518.

Bolivar, F. 1978. Construction and characterisation of new cloning vehicles, III. Derivatives of plasmid pBR322 carrying unique Eco RI sites for the selection of Eco RI generated recombinant molecules. *Gene* 4:121-136.

Busch, S. J., and P. Sassone-Corsi. 1990. Dimers, leucine zippers and DNA-binding domains. *Trends Genet.* 6:36-40.

Calame, K. L. 1989. Immunoglobulin gene transcription: molecular mechanisms. *Trends Genet.* 5:395-399.

Campbell, M. E. M., J. W. Palfreyman and C. M. Preston. 1984. Identification of herpes simplex virus DNA sequences which encode a transactivating polypeptide responsible for stimulation of immediate early transcription. *J. Mol. Biol.* 180:1-19.

Carbon, P., S. Murgo, J-P. Ebel, A. Krol, G. Tebb and I. W. Mattaj. 1987. A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II. *Cell* 51:71-79.

Chen, R., H. A. Ingraham, M. N. Treacy, V. R. Albert, L. Wilson and M. G. Rosenfeld. 1990. Autoregulation of pit-1 gene expression mediated by two cis-active promoter elements. *Nature* 346:583-586.

Chirgwin, J. M., A. E. Przybyla, R. J. McDonald and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources rich in ribonucleases. *Biochem.* 18:5294-5299.

Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore and P. A. Sharp. 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes Dev.* 2:1570-1581.

Cockerill, P. N., and S. P. Klinken. 1990. Octamer-binding proteins in diverse hemopoietic cells. *Mol. Cell. Biol.* 10:1293-1296.

Cox, P. M., S. M. Temperly, H. Kumar and C. R. Goding. 1988. A distinct octamer-binding protein present in malignant melanoma cells. *Nucleic Acids Res.* 16:11047-11056.

Currie, R. A., and R. G. Roeder. 1989. Identification of an octamer-binding site in the mouse kappa light-chain immunoglobulin enhancer. *Mol. Cell. Biol.* 9:4239-4247.

Dailey, L., S. B. Roberts and N. Heintz. 1988. Purification of the human histone H4 gene-specific transcription factors H4TF-1 and H4TF-2. *Genes Dev.* 2:1700-1712.

Dalton, S., and J. R. E. Wells. 1988. A gene-specific promoter element is required for optimal expression of the histone H1 gene in S-phase. *EMBO J.* 7:49-56.

Dalton, S., and J. R. E. Wells. 1988A. Maximal Binding Levels of an H1 Histone Gene-Specific Factor in S-Phase Correlate With Maximal H1 Gene Transcription. *Mol. Cell. Biol.* 8:4576-4578.

Davidson, I., E. Fromental, P. Augereau, A. Wildeman, M. Zenke and P. Chambon. 1986. Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. *Nature* 323:544-548.

Dawid, I. B. 1965. Deoxyribonucleic Acid in Amphibian Eggs. *J. Mol. Biol.* 12:581-599.

Dressler, G. E., and P. Gruss. 1988. Do multigene families regulate vertebrate development? *Trends Genet.* 4:214-219.

Dressler, G. E. 1989. An update on the vertebrate homeobox. *Trends Genet.* 5:129-130.

Dreyfus, M., N. Doyen and F. Rougeon. 1987. The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. EMBO J. 6:1685-1690.

Dumont, J. N. 1972. Oogenesis in X. laevis (Daudin). I. Stages of oocyte development in laboratory maintained animals. J. Morphol. 136:153-164.

Eaton, S., and K. L. Calame. 1987. Multiple DNA sequences are necessary for the function of an immunoglobulin heavy chain promoter. Proc. Natl. Acad. Sci. USA 84:7634-7638.

Elsholtz, H. P., V. R. Albert, M. N. Treacy and M. G. Rosenfeld. 1990. A two-base change in a POU factor-binding site switches pituitary-specific to lymphoid-specific gene expression. Genes Dev. 4:43-51.

Falkner, F., and H. Zachau. 1984. Correct transcription of an immunoglobulin k gene requires an upstream fragment containing conserved sequence elements. Nature 310:71-74.

Finney, M., G. Ruvkun and H. R. Horvitz. 1988. The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. Cell 55:757-769.

Fletcher, C., N. Heintz and R. G. Roeder. 1987.
Purification and Characterisation of OTF-1, a Transcription
Factor Regulating Cell Cycle Expression of a Human Histone
H2b Gene. Cell 51:773-781.

Gaffney, D. F., J. McLauchlin, J. L. Whitton and J. P.
Clements. 1985. A modular system for the assay of
transcription signals: the sequence TAATGARAT is required
for the herpes simplex virus immediate early gene
activation. Nucleic Acids Res. 13:7847-7863.

Garcia-Blanco, M. A., R. G. Clerc and P. A. Sharp. 1989.
The DNA-binding homeo domain of the Oct-2 protein. Genes
Dev. 3:739-745.

Gehring, W. J. 1987. Homeo Boxes in the Study of
Development. Science 236:1245-1251.

Gellman, I. H., and S. Silverstein. 1987. Dissection of
immediate-early gene promoters from herpes simplex virus:
sequences that respond to the virus transcriptional
activators. J. Virol. 61:3167-3172.

Gerster, T., P. Matthias, M. Thali, J. Jiricny and W.
Schaffner. 1987. Cell type-specificity elements of the
immunoglobulin heavy chain gene enhancer. EMBO J. 6:1323-
1330.

Gerster, T., and R. G. Roeder. 1988. A herpes virus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins. *Proc. Natl. Acad. Sci. USA* 85:6347-6351.

Gerster, T., C-M. Balmaceda and R. G. Roeder. 1990. The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J.* 9:1635-1643.

Gill, G., and M. Ptashne. 1988. Negative effect of the transcriptional activator GAL4. *Nature* 334:721-724.

Guiochon-Mantel, A., H. Loosfelt, P. Lescop, S. Sar, M. Atger, M. Perrot-Applanat and E. Milgrom. 1989. Mechanisms of nuclear localisation of the progesterone receptor: evidence for interaction between monomers. *Cell*:1147-1154.

Gumucio, D. L., K. L. Rood, T. A. Gray, M. F. Riordan, C. I. Sartar and F. S. Collins. 1988. Nuclear proteins that bind the human γ -globin gene promoter: alterations in binding produced by point mutations associated with hereditary persistence of fetal hemoglobin. *Mol. Cell. Biol.* 8:5310-5322.

Gusse, M., J. Ghysdael, G. Evan, T. Soussi and M. Mechali. 1989. Translocation of a store of maternal cytoplasmic C-myc protein into nuclei during early development. Mol. Cell. Biol. 9:5395-5403.

Hanahan, D. 1983. Studies on the transformation of Escherichia coli with plasmids. J. Mol. Biol. 166:557-580.

Harlow, E., and D. Lane. 1988. Antibodies: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Hatzopoulos, A. K., A. S. Stoykova, J. R. Erselius, M. Goulding, T. Neuman and P. Gruss. 1990. Structure and expression of the mouse Oct-2a and Oct-2b, two differentially spliced products of the same gene. Development 109:349-362.

He, X., M. N. Treacy, D. M. Simmons, H. A. Ingraham, L. W. Swanson and M. G. Rosenfeld. 1989. Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature 340:35-42.

Heindl, L. M., T. S. Weil and M. Perry. 1988. Promoter sequences required for transcription of Xenopus laevis histone genes in injected frog oocyte nuclei. Mol. Cell. Biol. 8:3676-3682.

Herr, W., R. A. Sturm, R. G. Clerc, L. M. Corcoran, D. Baltimore, P. A. Sharp, H. A. Ingraham, M. G. Rosenfeld, M. Finney, G. Ruvkun and H. R. Horvitz. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2 and Caenorhabditis elegans unc-86 gene products. 1988. Genes Dev. 2:1513-1516.

Hunt, T. 1989. Cytoplasmic anchoring proteins and the control of nuclear localisation. Cell 59:949-951.

Ingraham, H. A., R. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson and M. G. Rosenfeld. 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell 55:519-529.

Ito, M., A. Sharma, A. S. Lee and R. Maxson. 1989. Cell Cycle Regulation of H2b Histone Octamer DNA-Binding Activity in Chinese Hamster Lung Fibroblasts. Mol. Cell. Biol. 9:869-873.

Johnson, W. A., and J. Hirsh. 1990. Binding of a Drosophila POU-domain protein to a sequence element regulating gene expression in specific dopaminergic neurons. Nature 343:467-470.

Jones, N. C., P. W. J. Rigby and E. B. Ziff. 1988. Trans-acting protein factors and the control of eukaryotic transcription: lessons from studies on DNA tumour viruses. *Genes Dev.* 2:267-281.

Junker, S., S. Pederson, E. Schreiber and P. Matthias. 1990. Extinction of an immunoglobulin k promoter in cell hybrids is mediated by the octamer motif and correlates with suppression of Oct-2 expression. *Cell* 61:467-474.

Kamens, J., P. Richardson, G. Mosialos, R. Brent and T. Gilmore. 1990. Oncogenic transformation by vRel requires an amino-terminal activation domain. *Mol. Cell. Biol.* 10:2840-2847.

Karin, M., J-L. Castrillo and L. E. Theill. 1990. Growth hormone gene regulation: a paradigm for cell-type-specific gene activation. *Trends Genet.* 6:92-96.

Kessler, I., E. Schreiber, M. M. Muller, P. Matthias and W. Schaffner. 1989. Octamer transcription factors bind to two different sequence motifs of the immunoglobulin heavy chain promoter. *EMBO J.* 8:2001-2008.

Kemp, L. M., and D. S. Latchman. 1988. Differential regulation of octamer-containing cellular genes by herpes simplex virus virion protein Vmw65 is mediated by sequence differences in the octamer element. EMBO J. 7:4239-4244.

Kit, S., D. Dubbs, L. Piekaski and T. Hsu. 1963. Deletion of thymidine kinase activity from L cells resistant to bromodeoxyuridine. Exp. Cell. Res. 31:297-312.

Ko, H-S., P. Fast, W. McBride and L. M. Staudt. 1988. A Human Protein Specific for the Immunoglobulin Octamer DNA Motif Contains a Functional Homeobox Domain. Cell 55:135-144.

Kozack, M. 1989. The scanning model for translation: an update. J. Cell. Biol. 108:229-241.

Kristie, T. M., J. H. LeBowitz and P. A. Sharp. 1989. The octamer-binding proteins form multi-protein-DNA complexes with HSV α TIF regulatory protein. EMBO J. 8:4229-4238.

Krol, A., E. Lund and J. E. Dahlberg. 1985. The two embryonic U1 RNA genes of Xenopus laevis have both common and gene-specific transcription signals. EMBO J. 4:1529-1535.

Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.

LaBella, F., H. L. Sive, R. G. Roeder and N. Heintz. 1988. Cell-cycle regulation of a human histone H2b gene is mediated by the H2b subtype-specific consensus element. *Genes Dev.* 2:32-39.

Landolfi, N. F., J. D. Capra and P. W. Tucker. 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin VH promoter region sequences. *Nature* 323:548-551.

Landolfi, N. F., X-M. Yin, J. D. Capra and P. W. Tucker. 1988. A conserved heptamer upstream of the IgH promoter region octamer can be the site of a coordinate protein-DNA interaction. *Nucleic Acids Res.* 16:5503-5514.

Latchman, D. S., J. F. Partridge, J. K. Estridge and L. M. Kemp. 1989. The different competitive abilities of viral TAATGARAT elements and cellular octamer motifs, mediate the induction of viral immediate-early genes and the repression of the histone H2B gene in herpes simplex virus infected cells. *Nucleic Acids Res.* 17:8533-8542.

LeBowitz, J. H., T. Kobayashi, L. Staudt, D. Baltimore and P. A. Sharp. 1988. Octamer-binding proteins from B or HeLa cells stimulate transcription of the immunoglobulin heavy-chain promoter in vitro. *Genes Dev.* 2:1227-1237.

LeBowitz, J. H., R. G. Clerc, M. Brenowitz and P. A. Sharp. 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. *Genes Dev.* 3:1625-1638.

Lenardo, M., J. Pierce and D. Baltimore. 1987. Protein-binding sites in the Ig gene enhancers determine transcriptional activity and inducibility. *Science* 236:1573-1577.

Lenardo, M. J., and D. Baltimore. 1989. NF-kB: a pleiotropic mediator of inducible and tissue-specific gene control. *Cell* 58:227-229.

Lenardo, M. J., L. Staudt, P. Robbins, A. Kuang, R. C. Mulligan and D. Baltimore. 1989. Repression of the IgH enhancer in tetracarcinoma cells associated with a novel octamer factor. *Science* 243:544-546.

Levine, M., and T. Hoey. 1988. Homeobox Proteins as Sequence-Specific Transcription Factors. *Cell* 55:537-540.

Lund, E., and J. E. Dahlberg. 1987. Differential accumulation of U1 and U4 small nuclear RNAs during Xenopus development. *Genes Dev.* 1:39-46.

Mackem, S., and B. Roizman. 1982. Structural features of the virus α gene 4, 0, and 27 promoter-regulatory sequences which confer regulation on chimeric thymidine kinase genes. *J. Virol.* 44:939-949.

Mangalam, H. J., V. R. Albert, H. A. Ingraham, M. Kapiloff, L. Wilson, C. Nelson, H. Elsholtz and M. G. Rosenfeld. 1989. A pituitary POU domain protein, Pit-1, activates both growth hormone and prolactin promoters transcriptionally. *Genes Dev.* 3:946-958.

Maniatis, T., E. T. Fritsch and J. Sambrook. 1982. *Molecular cloning, a laboratory manual.* Cold Spring Harbor Laboratory, New York.

Maraden, H. S., M. E. M. Campbell, L. Haarr, M. C. Frame, D. S. Parriss, M. Murphy, R. G. Hope, M. T. Muller and C. M. Preston. 1987. The 65000-Mr DNA binding and virion trans-inducing proteins of herpes simplex virus type 1. *J. Virol.* 61:2428-2437.

- Mattaï, I. W., S. Lienhard, J. Jiricny and E. M. DeRobertis. 1985. An enhancer-like sequence within the Xenopus U2 gene promoter facilitates the formation of stable transcription complexes. *Nature* 316:163-167.
- McCormick, A., H. Brady, L. E. Theill and M. Karin. 1990. Regulation of the pituitary-specific gene GHF-1 by cell-autonomous and environmental cues. *Nature* 345:829-832.
- McKnight, J. L. C., T. M. Kristie and B. Roizman. 1987. Binding of the virion protein mediating α gene induction in herpes simplex virus 1 - infected cells to its cis site requires cellular proteins. *Proc. Natl. Acad. Sci. USA* 84:7061-7065.
- Messing, J., R. Crea and P. H. Seeburg. 1988. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 9:309-321.
- Mitchell, P. J., and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245:371-378.
- Moorman, A. F. M., P. A. J. DeBoer, R. T. M. DeLaaf and O. H. J. Destree. 1982. Primary structure of the histone H2A and H2B genes and their flanking sequences in a minor histone gene cluster of Xenopus laevis. *FEBS Letters* 144:235-241.

Muller, M. M., S. Ruppert, W. Schaffner and P. Matthias. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature* 336:544-551.

Muller-Immergluck, M. M., W. Schaffner and P. Matthias. 1990. Transcription factor Oct-2A contains functionally redundant activating domains and works selectively from a promoter but not from a remote enhancer position in non-lymphoid (HeLa) cells. *EMBO J.* 9:1625-1634.

Murphy, S., C. DiLiegro and M. Melli. 1987. The in vitro transcription of the 7SK RNA gene by RNA pol III is dependent only on the presence of an upstream promoter. *Cell* 51:81-87.

Murphy, S., A. Pierani, C. Scheidereit, M. Melli and R. G. Roeder. 1989. Purified Octamer Binding Transcription Factors Stimulate RNA Polymerase III- Mediated Transcription of the 7SK RNA Gene. *Cell* 59:1071-1080.

Nelson, C., V. R. Albert, H. P. Elsholtz, L. I-W. Lu and M. G. Rosenfeld. 1988. Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. *Science* 239:1400-1405.

Nieuwkoop, P. D., and J. Faber. 1956. Normal table of Xenopus laevis (Daudin). North-Holland Publishing Co., Amsterdam.

O'Hare, P., and C. R. Goding. 1988. Herpes simplex virus regulatory elements and the immunoglobulin octamer domain bind a common factor and are both targets for virion transactivation. *Cell* 52:435-445.

O'Hare, P., C. R. Goding and A. Haigh. 1988. Direct combinatorial interaction between a herpes simplex virus regulatory protein and a cellular octamer-binding factor mediates specific induction of virus immediate-early gene expression. *EMBO J.* 7:4231-4238.

Okamoto, K., H. Okazama, A. Okuda, M. Sakai, M. Muramatsu and H. Hamada. 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* 60:461-472.

Old, R. W., and H. R. Woodland. 1984. Histone genes: not so simple after all. *Cell* 38:624-626.

Old, R. W., S. A. Sheikh, A. Chambers, C. A. Newton, A. Mohammed and T. C. Aldridge. 1985. Individual Xenopus histone genes are replication-independent in oocytes and replication-dependent in Xenopus or mouse somatic cells. Nucleic Acids Res. 13:7341-7358.

Old, R. W., G. E. Sweeney and A. R. Brooks. 1988. Activity of a cloned Xenopus albumin gene in the homologous frog oocyte system. Biochem. Biophys. Acta 951:220-225.

Parry, H. D., D. Scherly and I. W. Mattaj. 1989. 'Snurpogenesis': the transcription and assembly of U snRNP components. Trends Biochem. Sci. 14:15-19.

Perry, M., G. H. Thomsen and R. G. Roeder. 1985. Genomic Organisation and Nucleotide Sequence of Two Distinct Histone Gene Clusters from Xenopus laevis. J. Mol. Biol. 185:479-499.

Perry, M., G. H. Thomsen and R. G. Roeder. 1986. Major transitions in histone gene expression do not occur during development in Xenopus laevis. Dev. Biol. 116:532-538.

Petryniak, B., L. M. Staudt, C. E. Postema, W. T. McCormack and C. B. Thompson. 1990. Characterisation of chicken octamer-binding proteins demonstrates that POU domain-containing homeobox transcription factors have been highly conserved during vertebrate evolution. *Proc. Natl. Acad. Sci. USA* 87:1099-1103.

Picard, D., and K. R. Yamamoto. 1987. Two signals mediate hormone-dependent nuclear localisation of the glucocorticoid receptor. *EMBO J.* 6:3333-3340.

Poellinger, L., and R. G. Roeder. 1989. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. *Mol. Cell. Biol.* 9:747-756.

Prentki, P., F. Karch, S. Iida and J. Meyer. 1981. The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. *Gene* 14:289-299.

Preston, C. M., M. C. Frame and M. E. M. Campbell. 1988. A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. *Cell* 52:425-434.

Pruijn, G. J. M., W. van Driel, R. T. van Miltenburg and P. C. van der Vliet. 1987. Promoter and enhancer elements containing a conserved sequence motif are recognised by nuclear factor III, a protein stimulating adenovirus DNA replication. EMBO J. 6:3771-3778.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. Nature 335:683-689.

Pudney, M., M. G. R. Varma and C. J. Leake. 1973. Establishment of a cell line (XTC-2) from the South African clawed toad, Xenopus laevis. Experientia 29:466-467.

Roeder, R. G. 1974. Multiple forms of deoxyribonucleic acid-dependent ribonucleic acid polymerase in Xenopus laevis. J. Biol. Chem. 249:249-256.

Rosales, R., M. Vigneron, M. Macchi, I. Davidson, J. H. Xiao and P. Chambon. 1987. In vitro binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. EMBO J. 10:3015-3025.

Rosenberg, A. H., B. N. Lade, D-S. Chui, S-W. Lin, J. J. Dunn and F. W. Studier. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene 56:125-135.

Rosner, M. H., M. A. Viganò, K. Ozato, P. M. Timmons, F. Poirier, P. W. J. Rigby and L. M. Staudt. 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* 345:686-692.

Roth, S., D. Stein and C. Nusslein-Volhard. 1989. A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the Drosophila embryo. *Cell* 59:1189-1202.

Rushlow, C. A., K. Han, J. L. Manley and M. Levine. 1989. The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. *Cell* 59:1165-1177.

Sadowski, I., J. Ma, S. Triezenberg and M. Ptashne. 1988. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* 335:563-564.

Sanger, F., S. Nicklen and A. R. Coulson. 1977. DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

- Scheidereit, C., A. Heguy and R. G. Roeder. 1987. Identification and Purification of a Human Lymphoid-Specific Octamer-Binding Protein (OTF-2) That Activates Transcription of an Immunoglobulin Promoter In Vitro. *Cell* 51:783-793.
- Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C-G. Balmaceda, R. A. Currie and R. G. Roeder. 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. *Nature* 336:551-557.
- Scholer, H. R., A. K. Hatzopoulos, R. Balling, N. Suzuki and P. Gruss. 1989. A family of octamer-specific proteins present during embryogenesis: evidence for germline-specific expression of an Oct factor. *EMBO J.* 8:2543-2550.
- Scholer, H. R., R. Balling, A. K. Hatzopoulos, N. Suzuki and P. Gruss. 1989A. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *EMBO J.* 8:2551-2557.
- Scholer, H. R., S. Ruppert, N. Suzuki, K. Chowdhury and P. Gruss. 1990. New type of POU domain in germ line-specific protein Oct-4. *Nature* 344:435-439.

Schreiber, E., P. Matthias, M. M. Muller and W. Schaffner. 1988. Identification of a novel lymphoid specific octamer binding protein (OTF-28) by proteolytic clipping band shift assay (PCBA). EMBO J. 7:4221-4229.

Schumperli, D. 1988. Multilevel regulation of replication-dependent histone genes. Trends Genet. 4:187-191.

Short, J. M., J. M. Fernandez, J. A. Sorge and W. D. Huse. 1988. Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. Nucleic Acids Res. 16:7583-7601.

Simmons, D. M., J. W. Voss, H. A. Ingraham, J. M. Holloway, R. S. Broide, M. G. Rosenfeld and L. W. Swanson. 1990. Pituitary cell phenotypes involve cell-specific Pit-1 mRNA translation and synergistic interactions with other classes of transcription factors. Genes Dev. 4:695-711.

Singh, H., R. Sen, D. Baltimore and P. A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature 319:154-158.

Sive, H. L., and R. G. Roeder. 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin, and U2 small nuclear RNA (snRNA) genes. Proc. Natl. Acad. Sci. USA 83:6382-6386.

Sive, H. L., N. Heintz and R. G. Roeder. 1986. Multiple sequence elements are required for maximal in vitro transcription of a human histone H2B gene. Mol. Cell. Biol. 6:3329-3340.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature 323:640-643.

Staudt, L. M., R. G. Clerc, H. Singh, J. H. LeBowitz, P. A. Sharp and D. Baltimore. 1988. Cloning of a Lymphoid-Specific cDNA Encoding the Regulatory Octamer DNA Motif. Science 241:577-580.

Stern, S., M. Tanaka and W. Herr. 1989. The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature* 341:624-630.

Steward, R. 1989. Relocalisation of the dorsal protein from the cytoplasm to the nucleus correlates with its function. *Cell* 59:1179-1188.

Stringer, K. F., C. J. Ingles and J. Greenblatt. 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* 345:783-786.

Studier, F. W., and B. A. Moffatt. 1986. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. *J. Mol. Biol.* 189:113-130.

Sturm, R., T. Baumruker, B. R. Franza Jr. and W. Herr. 1987. A 100-kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. *Genes Dev.* 1:1147-1160.

Sturm, R. A., S. Dalton and J. R. E. Wells. 1988.

Conservation of histone H2A/H2B intergene regions: a role for the H2B specific element in divergent transcription. Nucleic Acids Res. 16:8571-8586.

Sturm, R. A., G. Das and W. Herr. 1988A. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. 2:1582-1599.

Sturm, R. A., and W. Herr. 1988. The POU domain is a bipartite DNA-binding structure. Nature 336:601-604.

Takimoto, M., J. P. Quinn, A. R. Farina, L. M. Staudt and D. Levens. 1989. fos/jun and octamer-binding protein interact with a common site in the negative element of the human c-myc gene. J. Biol. Chem. 264:8992-8999.

Tanaka, M., U. Grossniklaus, W. Herr and N. Hernandez. 1988. Activation of the U2 snRNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. Genes Dev. 2:1764-1778.

Tanaka, M., and W. Herr. 1990. Differential Transcriptional Activation by Oct-1 and Oct-2: Interdependent Activation Domains Induce Oct-2 Phosphorylation. Cell 60:375-386.

Thomas, P. S. 1980. Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77:5201-5205.

Wigg, A. J., and D. Sherratt. 1980. Trans-complementable copy-number mutants of plasmid ColE1. nature 283:216-218.

Verrijzer, C. P., A. J. Kal and P. C. van der Vliet. 1990. The DNA binding domain (POU domain) of transcription factor oct-1 suffices for the stimulation of DNA replication. EMBO J. 9:1883-1888.

Wells, D. E. 1986. Compilation analysis of histones and histone genes. Nucl. Acids Res. 14:r119-r149.

Wirth, T., L. Staudt and D. Baltimore. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature 329:174-178.

Woodland, H. 1982. The translational control phase of early development. Biosci. Rep. 2:471-491.

Woodland, H. R., R. W. Old, E. A. Sturges, J. E. M. Ballantine, T. C. Aldridge and P. C. Turner. 1983. The strategy of histone gene expression in the development of Xenopus. Current problems in germ cell differentiation. Symposium of the British Society for Developmental Biology. 353-376.

Woodland, H. R., J. R. Warmington, J. E. M. Ballantine and P. C. Turner. 1984. Are there major developmentally regulated H4 gene classes in Xenopus? Nucleic Acids Res. 12:4939-4958.

Wright, C. V. E., K. W. Y. Cho, G. Oliver and E. M. DeRobertis. 1989. Vertebrate homeodomain proteins: families of region-specific transcription factors. Trends Biochem. Sci. 14:52-56.

Yu, H., B. Porton, L. Shen and L. A. Eckhardt. 1989. Role of the octamer motif in hybrid cell extinction of immunoglobulin gene expression: extinction is dominant in a 2 cell enhancer system. Cell 58:441-448.

Zhong, R., R. G. Roeder and N. Heintz. 1983. The primary structure and expression of four cloned human histone genes. Nucleic Acids Res. 11:7409-7425.

Ziff, E. B. 1990. Transcription factors: a new family gathers at the cAMP response site. Trends Genet. 6:69-72.

Zinn, K., D, DiMaio and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human B-interferon gene. Cell 34:865-879.